

ON THE SYLVESTER MATRIX EQUATION OVER QUATERNIONS

VLADIMIR BOLOTNIKOV

ABSTRACT. The Sylvester equation $AX - XB = C$ is considered in the setting of quaternion matrices. Conditions that are necessary and sufficient for the existence of a unique solution are well-known. We study the complementary case where the equation either has infinitely many solutions or does not have solutions at all. Special attention is given to the case where A and B are respectively, lower and upper triangular two-diagonal matrices (in particular, if A and B are Jordan blocks).

1. INTRODUCTION

We start with the complex case: given complex matrices $A \in \mathbb{C}^{n \times n}$, $B \in \mathbb{C}^{m \times m}$, $C \in \mathbb{C}^{n \times m}$, the Sylvester equation

$$AX - XB = C \quad (1.1)$$

has a unique solution $X = [x_{ij}] \in \mathbb{C}^{n \times m}$ if and only if the spectrums $\sigma(A)$ and $\sigma(B)$ are disjoint. The result was established by Sylvester [17] via representing (1.1) in the equivalent form $G\mathbf{x} = \mathbf{c}$ where $G = A \otimes I_m - I_n \otimes B^\top$ and where \mathbf{x} , \mathbf{c} are the columns constructed from the entries of X and C by

$$\mathbf{x} = \text{Col}_{1 \leq i \leq n} (\text{Col}_{1 \leq i \leq m} x_{ij}), \quad \mathbf{c} = \text{Col}_{1 \leq i \leq n} (\text{Col}_{1 \leq i \leq m} c_{ij}). \quad (1.2)$$

Thus, the equation (1.1) has a unique solution (for any $C \in \mathbb{C}^{n \times m}$) if and only if the matrix G is invertible (equivalently, $\sigma(A) \cap \sigma(B) = \emptyset$), in which case the unique solution X is recovered from the column $G^{-1}\mathbf{c}$. The infinite-dimensional extension of the Sylvester's theorem as well as the integral formula for the unique solution X of (1.1) is due to Rosenblum [14]. Some other explicit formulas for $X \in \mathbb{C}^{n \times m}$ (in terms of A , B and C rather than their entries) can be found in the survey [11]. If $\sigma(A) \cap \sigma(B) \neq \emptyset$, then (1.1) has a solution if and only if $\text{rank} G = \text{rank} [G \ \mathbf{c}]$, while the homogeneous equation $AX = XB$ has d linearly independent solutions where the integer $d = \dim \text{Nul}(G)$ can be expressed in terms of invariant factors of matrix pencils A and B [5, 7]).

Let $\mathcal{J}_n(\alpha)$ denote the $n \times n$ lower triangular Jordan block with α on the main diagonal:

$$\mathcal{J}_n(\alpha) = \alpha I_n + F_n, \quad \text{where} \quad F_n = [\delta_{i,j+1}]_{i,j=1}^n \quad (1.3)$$

and where $\delta_{i,j}$ is the Kronecker symbol. If A and B are Jordan blocks $A = \mathcal{J}_n(\alpha)$ and $B = \mathcal{J}_m^\top(\beta)$, the equation (1.1) simplifies to

$$(\alpha - \beta)X = C - F_n X + X F_m^\top. \quad (1.4)$$

If $\alpha \neq \beta$, we iterate the latter equality $n + m - 1$ times and arrive at

$$X = \sum_{k=0}^{n+m-2} (\alpha - \beta)^{-k-1} \sum_{i+j=k} (-1)^{k+i} \binom{k}{i} F_n^i C (F_m^\top)^j. \quad (1.5)$$

The formula (1.5) appeared in [13] and in a slightly different form, in [15]. If $\alpha = \beta$ and $n \geq m$ (this case was considered in [13]), the equation (1.4) can be written entry-wise as follows:

$$\begin{aligned} 0 &= c_{1,1}, \quad x_{i,1} = c_{i+1,1}, \quad x_{1,j} = -c_{1,j+1}, \\ x_{i,j+1} - x_{i+1,j} &= c_{i+1,j+1} \quad \text{for } 1 \leq i < n; \quad 1 \leq j < m. \end{aligned} \quad (1.6)$$

Then one can see that if the system (1.6) has a solution, then necessarily

$$\sum_{i=1}^{k-1} c_{i,k-i} = 0 \quad \text{for } k = 2, \dots, n. \quad (1.7)$$

On the other hand if conditions (1.7) are met, then for any choice of fixed $x_{n,1}, \dots, x_{n,m}$, the system (1.5) has a unique solution. In other words, the bottom row in X serves as a free and independent parameter in the parametrization of all solutions X of the Sylvester equation (1.4).

In general, one may reduce given matrices A and B to their lower and upper Jordan forms (with say, k and ℓ Jordan cells, respectively) subsequently splitting the equation (1.1) to $k\ell$ Sylvester equations of the form (1.4).

In this paper, we will focus on the equation (1.1) over quaternions:

$$AX - XB = C, \quad \text{where } A \in \mathbb{H}^{n \times n}, B \in \mathbb{H}^{m \times m}, C \in \mathbb{H}^{n \times m}. \quad (1.8)$$

Since multiplication in \mathbb{H} is non-commutative, the equation (1.8) is not completely trivial even in the scalar case. The following result back to Hamilton (see e.g., [18, p. 123]).

Theorem 1.1. *Given $\alpha, \beta, c \in \mathbb{H}$, the equation $\alpha x - x\beta = c$ has a unique solution if and only if*

$$P_{\alpha,\beta} := |\alpha|^2 - (\alpha + \bar{\alpha})\beta + \beta^2 \neq 0 \quad (1.9)$$

and this unique solution equals $x = (\bar{\alpha}c - c\beta)P_{\alpha,\beta}^{-1}$.

Taking the advantage of complex representations for quaternion matrices [12], it is always possible to reduce (1.8) to certain complex Sylvester equation producing in particular, the uniqueness criterion: *the equation (1.8) has a unique solution if and only if the right spectrums of A and B are disjoint*; see [9, 16] and Theorem 3.1 below. However, further results obtained on this way and briefly surveyed in Section 3 refer, to some extent, to complex representations of matrices A and B rather than the matrices themselves; see e.g., formula (3.3) for the unique solution.

Our contribution here are several explicit formulas for the unique solution given exclusively in terms of the original matrices A, B, C in the cases where (1) A and B are Jordan blocks (Theorem 3.2 presents the quaternion analog of formula (1.5)), (2) A is two-diagonal (Theorems 3.3 and 3.4), and (3) A is lower-triangular (Theorem 3.8). Making use of canonical Jordan forms for A and B (see [20] and Theorem 2.1 below) one can reduce the general case to the one where A and B are Jordan cells.

The core of the paper is the study of the singular case (the right spectrums of A and B are not disjoint). Special attention is given to the case where A and B are two-diagonal matrices (see formulas (2.21)). As will be explained in Section 2.4, Sylvester equations with A and B of this form arise in the context of polynomial interpolation over quaternions and the results on such special Sylvester equations are needed to explicitly describe quasi-ideals in the ring of quaternion polynomials. In Section 4 we present necessary and sufficient conditions for the equation to have a solution (Theorem 4.5) which become more transparent if A and B are Jordan blocks (Theorem 4.7 presents the analog of conditions (1.7) from [13]). Also in Section 4, we present an algorithm for constructing a solution to the (solvable) equation (1.8). In Section 5, we parametrize the solution set of the homogeneous equation $AX - XB = 0$; the parametrization contains $\min(m, n)$ free independent parameters, each one of which varies in a two-dimensional real subspace of \mathbb{H} . In case $A = \mathcal{I}_n(\alpha)$ and $B = \mathcal{I}_m(\beta)$ are Jordan blocks, the general solution of the homogeneous equation is a “triangular” Hankel matrix all entries of which satisfy the homogeneous scalar Sylvester equation $\alpha x - x\beta = 0$ (Corollary 5.2).

2. PRELIMINARIES

In this section we collect basic facts to make presentation self-contained. We first fix notation and terminology. By \mathbb{H} we denote the skew field of quaternions $\alpha = x_0 + \mathbf{i}x_1 + \mathbf{j}x_2 + \mathbf{k}x_3$ where $x_0, x_1, x_2, x_3 \in \mathbb{R}$ and where $\mathbf{i}, \mathbf{j}, \mathbf{k}$ are the imaginary units commuting with \mathbb{R} and satisfying $\mathbf{i}^2 = \mathbf{j}^2 = \mathbf{k}^2 = \mathbf{ijk} = -1$. For $\alpha \in \mathbb{H}$ as above, its real and imaginary parts, the quaternion conjugate and the absolute value are defined as $\text{Re}(\alpha) = x_0$, $\text{Im}(\alpha) = \mathbf{i}x_1 + \mathbf{j}x_2 + \mathbf{k}x_3$, $\bar{\alpha} = \text{Re}(\alpha) - \text{Im}(\alpha)$ and $|\alpha|^2 = \alpha\bar{\alpha} = |\text{Re}(\alpha)|^2 + |\text{Im}(\alpha)|^2$, respectively. Two quaternions α and β are called *equivalent* (conjugate to each other) if $\alpha = h^{-1}\beta h$ for some nonzero $h \in \mathbb{H}$; in notation, $\alpha \sim \beta$. It turns out (see e.g., [4]) that

$$\alpha \sim \beta \quad \text{if and only if} \quad \text{Re}(\alpha) = \text{Re}(\beta) \text{ and } |\alpha| = |\beta|, \quad (2.1)$$

so that the *conjugacy class* of a given $\alpha \in \mathbb{H}$ form a 2-sphere (of radius $|\text{Im}(\alpha)|$ around $\text{Re}(\alpha)$) which will be denoted by $[\alpha]$. It is clear that $[\alpha] = \{\alpha\}$ if and only if $\alpha \in \mathbb{R}$.

A finite ordered collection $\boldsymbol{\alpha} = (\alpha_1, \dots, \alpha_n)$ will be called a *spherical chain* (of length n) if

$$\alpha_1 \sim \alpha_2 \sim \dots \sim \alpha_k \quad \text{and} \quad \alpha_{j+1} \neq \bar{\alpha}_j \quad \text{for} \quad j = 1, \dots, n-1. \quad (2.2)$$

The latter notion is essentially non-commutative: a spherical chain $\boldsymbol{\alpha} = (\alpha_1, \dots, \alpha_n)$ consisting of commuting elements necessarily belongs to the set $\{\alpha_1, \bar{\alpha}_1\}$ which together with inequality in (2.2) implies that all elements in $\boldsymbol{\alpha}$ are the same:

$$\boldsymbol{\alpha} = (\alpha, \alpha, \dots, \alpha), \quad \alpha \in \mathbb{H}. \quad (2.3)$$

2.1. Quaternion matrices. We denote by $\mathbb{H}^{n \times m}$ the space of $n \times m$ matrices with quaternion entries. The definitions of the transpose matrix A^\top , the quaternion-conjugate matrix \bar{A} and the adjoint matrix A^* are the same as in the complex case.

An element $\alpha \in \mathbb{H}$ is called a (right) eigenvalue of the matrix $A \in \mathbb{H}^{n \times n}$ if $A\mathbf{x} = \mathbf{x}\alpha$ for some nonzero $\mathbf{x} \in \mathbb{H}^{n \times 1}$. In this case, for any $\beta = h^{-1}\alpha h \sim \alpha$ we

also have $A\mathbf{x}h = \mathbf{x}hh^{-1}\alpha h = \mathbf{x}h\beta$ and hence, any element in the conjugacy class $[\alpha]$ is a right eigenvalue of A . Therefore, the right spectrum $\sigma_{\mathbf{r}}(A)$ is the union of disjoint conjugacy classes (some of which may be real singletons). The studies of right eigenvalues and canonical forms for quaternion matrices were carried out in [4, 12, 20]. In particular, it was shown in [20] that any square quaternion matrix is similar to a complex matrix in Jordan form.

Theorem 2.1. *For every $A \in \mathbb{H}^{n \times n}$, there is an invertible $S \in \mathbb{H}^{n \times n}$ such that $S^{-1}AS = \bigoplus_{i=1}^k \mathcal{J}_{n_i}(\alpha_i)$, where $\alpha_i \in \mathbb{C}$ and $\text{Im}\alpha_i \geq 0$ for $i = 1, \dots, k$.*

Making use of Theorem 2.1 one can reduce A and B to their lower and upper Jordan forms

$$S^{-1}AS = \bigoplus_{i=1}^k \mathcal{J}_{n_i}(\alpha_i), \quad T^{-1}BT = \bigoplus_{j=1}^{\ell} \mathcal{J}_{m_j}^{\top}(\beta_j)$$

and conformally decompose $S^{-1}CT = [C_{ij}]$ and $S^{-1}XT = [X_{ij}]$ to see that (1.8) splits into $k\ell$ Sylvester equations

$$\mathcal{J}_{n_i}(\alpha_i)X_{ij} - X_{ij}\mathcal{J}_{m_j}^{\top}(\beta_j) = C_{ij}. \quad (2.4)$$

This reduction suggests to study the “basic” case where A and B are Jordan blocks, i.e., to get the quaternion analogs of the formula (1.5) and conditions (1.7). The deficiency of this approach is that most of explicit formulas will rely on similarity matrices S and T .

2.2. Complex representations. Since each $\alpha \in \mathbb{H}$ admits a unique representation of the form $\alpha = \alpha_1 + \alpha_2\mathbf{j}$ with $\alpha_1, \alpha_2 \in \mathbb{C}$, any matrix $A \in \mathbb{H}^{n \times m}$ can be written uniquely as $A = A_1 + A_2\mathbf{j}$ with $A_1, A_2 \in \mathbb{C}^{n \times m}$. The map

$$A = A_1 + A_2\mathbf{j} \mapsto \varphi(A) = \begin{bmatrix} A_1 & A_2 \\ -\overline{A_2} & \overline{A_1} \end{bmatrix} \quad (2.5)$$

that associates to each quaternion matrix its *complex representation* was introduced in [12]. It is additive and multiplicative in the sense that

$$\varphi(A + B) = \varphi(A) + \varphi(B) \quad \text{and} \quad \varphi(AB) = \varphi(A)\varphi(B) \quad (2.6)$$

for rectangular quaternion matrices of appropriate sizes. It was shown in [12] that if $\lambda \in \mathbb{C}$ is an eigenvalue of $\phi(A)$, then any element from the 2-sphere $[\lambda] \subset \mathbb{H}$ is a right eigenvalue of A and that all right eigenvalues of A arise in this way (this fact implies that the right spectrum of an $n \times n$ matrix is the union of at most n disjoint conjugacy classes). Therefore, for matrices $A \in \mathbb{H}^{n \times n}$ and $B \in \mathbb{H}^{m \times m}$, the following conditions are equivalent:

$$\sigma(\varphi(A)) \cap \sigma(\varphi(B)) = \emptyset \iff \sigma_{\mathbf{r}}(A) \cap \sigma_{\mathbf{r}}(B) = \emptyset. \quad (2.7)$$

Observing that $\mathbf{j}D = \overline{D}\mathbf{j}$ for any complex matrix D we define the map

$$\begin{aligned} \psi : Y = \begin{bmatrix} Y_{11} & Y_{12} \\ Y_{21} & Y_{22} \end{bmatrix} &\mapsto \frac{1}{2} \begin{bmatrix} I_n & -\mathbf{j}I_n \end{bmatrix} Y \begin{bmatrix} I_m \\ \mathbf{j}I_m \end{bmatrix} \\ &= \frac{Y_{11} + \overline{Y}_{22}}{2} + \frac{Y_{12} - \overline{Y}_{21}}{2} \mathbf{j} \end{aligned} \quad (2.8)$$

assigning to each matrix $Y \in \mathbb{C}^{2n \times 2m}$ a quaternion matrix constructed from the blocks $Y_{ij} \in \mathbb{C}^{n \times m}$. It is easily verified that ψ is the left inverse of ϕ :

$$\psi(\phi(A)) = A \quad \text{for any } A \in \mathbb{H}^{n \times m}. \quad (2.9)$$

Besides, ψ is additive and, although not multiplicative, the equalities

$$\begin{aligned} \psi(\phi(A)Y) &= \psi(\phi(A))\psi(Y) = A\psi(Y), \\ \psi(Y\phi(B)) &= \psi(Y)\psi(\phi(B)) = \psi(Y)B \end{aligned} \quad (2.10)$$

hold for any $A \in \mathbb{H}^{n \times m}$, $B \in \mathbb{H}^{p \times q}$ and $Y \in \mathbb{C}^{2m \times 2p}$.

2.3. Quaternion polynomials. Let $\mathbb{H}[z]$ denote the ring of polynomials in one formal variable z which commutes with quaternionic coefficients. The ring operations in $\mathbb{H}[z]$ are defined as in the commutative case, but as multiplication in \mathbb{H} is not commutative, multiplication in $\mathbb{H}[z]$ is not commutative as well. For any $\alpha \in \mathbb{H}$, we define left and right evaluation of f at α by

$$f^{e_\ell}(\alpha) = \sum_{j=0}^k \alpha^j f_j \quad \text{and} \quad f^{e_r}(\alpha) = \sum_{j=0}^k f_j \alpha^j \quad \text{if} \quad f(z) = \sum_{j=0}^k z^j f_j. \quad (2.11)$$

The formulas make sense for matrix valued polynomials $f \in \mathbb{H}^{n \times m}[z]$ and extend to square matrices by letting

$$f^{e_\ell}(A) = \sum_{j=0}^k A^j f_j, \quad f^{e_r}(B) = \sum_{j=0}^k f_j B^j \quad \text{if} \quad A \in \mathbb{H}^{n \times n}, B \in \mathbb{H}^{n \times n}. \quad (2.12)$$

An element $\alpha \in \mathbb{H}$ is called a left (right) zero of f if $f^{e_\ell}(\alpha) = 0$ (respectively, $f^{e_r}(\alpha) = 0$). For polynomials with real coefficients, left and right evaluations (and therefore, the notions of left and right zeros) coincide. The *characteristic polynomial* of a non-real conjugacy class $[\alpha] \subset \mathbb{H}$ is defined by

$$\mathcal{X}_{[\alpha]}(z) = (z - \alpha)(z - \bar{\alpha}) = z^2 - z(\alpha + \bar{\alpha}) + |\alpha|^2; \quad (2.13)$$

it follows from characterization (2.1) that formula (2.13) does not depend on the choice of $\alpha \in [\alpha]$. Since $\mathcal{X}_{[\alpha]}$ is the polynomial of the minimally possible degree such that its zero set (left and right, as $\mathcal{X} \in \mathbb{R}[z]$) coincides with $[\alpha]$, it is also called the *minimal polynomial* of $[\alpha]$. Observe that the matrix $\mathcal{X}_{[\alpha]}(B)$ is invertible if and only if $[\alpha] \cap \sigma_r(B) = \emptyset$.

Since the division algorithm holds in $\mathbb{H}[z]$ on either side (see e.g., [19]), any (left or right) ideal in $\mathbb{H}[z]$ is principal. We will use notation $\langle h \rangle_r$ and $\langle h \rangle_\ell$ for respectively the right and the left ideal generated by h . An ideal is maximal if and only if it is generated by the polynomial

$$\rho_\alpha(z) = z - \alpha, \quad \alpha \in \mathbb{H} \quad (2.14)$$

and it follows from respective (left and right) division algorithms that

$$f \in \langle \rho_\alpha \rangle_r \Leftrightarrow f^{e_\ell}(\alpha) = 0, \quad f \in \langle \rho_\alpha \rangle_\ell \Leftrightarrow f^{e_r}(\alpha) = 0. \quad (2.15)$$

A left (right) ideal is called *irreducible* if it is not contained properly in two distinct left (right) ideals which occurs if and only if it is generated by a polynomial p of the form $p = \rho_{\alpha_1} \rho_{\alpha_2} \cdots \rho_{\alpha_n}$ for some spherical chain $\alpha = (\alpha_1, \dots, \alpha_n)$ (see e.g., [2]).

2.4. Sylvester equations and interpolation by polynomials. The equation (1.8) arises naturally in the context of the following interpolation problem: *given polynomials $p, \tilde{p}, g, \tilde{g}$, find a polynomial $f \in \mathbb{H}[z]$ such that*

$$f - g \in \langle p \rangle_{\mathbf{r}} \quad \text{and} \quad f - \tilde{g} \in \langle \tilde{p} \rangle_{\mathbf{l}}. \quad (2.16)$$

If $p = \rho_{\alpha}$ and $q = \rho_{\beta}$, then due to equivalences (2.15), conditions (2.16) can be written in terms of left and right evaluations as

$$f^{e_{\mathbf{l}}}(\alpha) = \gamma \quad \text{and} \quad f^{e_{\mathbf{r}}}(\beta) = \delta \quad (2.17)$$

where $\gamma = h^{e_{\mathbf{l}}}(\alpha)$ and $\delta = g^{e_{\mathbf{r}}}(\beta)$. The following result appears in [1].

Theorem 2.2. *There is a polynomial $f \in \mathbb{H}[z]$ satisfying conditions (2.17) if and only if the Sylvester equation*

$$\alpha x - x \beta = \gamma - \delta \quad (2.18)$$

has a solution. If this is the case, all polynomials $f \in \mathbb{H}[z]$ satisfying conditions (2.17) are parametrized by the formula

$$f = \gamma + \rho_{\alpha} x + \rho_{\alpha} h \rho_{\beta} \quad (= \delta + x \rho_{\beta} + \rho_{\alpha} h \rho_{\beta}) \quad (2.19)$$

when x runs through the solution set of (2.18) and h varies in $\mathbb{H}[z]$.

Note that letting $h \equiv 0$ in (2.19) gives all linear solutions to the problem (2.17) while letting $\gamma = \delta = 0$ leads us to the explicit description of the intersection $\langle \rho_{\beta} \rangle_{\mathbf{l}} \cap \langle \rho_{\alpha} \rangle_{\mathbf{r}}$ of two maximal ideals in $\mathbb{H}[z]$.

Let us now assume that the ideals $\langle p \rangle_{\mathbf{r}}$ and $\langle q \rangle_{\mathbf{l}}$ in (2.16) are irreducible, i.e., that p and q are of the form

$$p = \rho_{\alpha_1} \rho_{\alpha_2} \cdots \rho_{\alpha_n}, \quad q = \rho_{\beta_m} \rho_{\beta_{m-1}} \cdots \rho_{\beta_1} \quad (2.20)$$

for some spherical chains where $\alpha = (\alpha_1, \dots, \alpha_n)$ and $\beta = (\beta_1, \dots, \beta_m)$. Let us introduce the matrices

$$\mathcal{J}_{\alpha} = \begin{bmatrix} \alpha_1 & 0 & \cdots & 0 \\ 1 & \alpha_2 & 0 & \\ 0 & 1 & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & 1 & \alpha_n \end{bmatrix}, \quad \mathcal{J}_{\beta} = \begin{bmatrix} \beta_1 & 0 & \cdots & 0 \\ 1 & \beta_2 & 0 & \\ 0 & 1 & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & 1 & \beta_m \end{bmatrix} \quad (2.21)$$

and let $E_n \in \mathbb{H}^{n \times 1}$ be the column with the top entry equal one and all other entries equal zero.

Theorem 2.3. *Let p and q be as in (2.20) and, given polynomials $g(z) = \sum g_j z^j$ and $\tilde{g}(z) = \sum \tilde{g}_k z^k$ with $\deg g < \deg p$ and $\deg \tilde{g} < \deg \tilde{p}$, let*

$$C = \sum_j \mathcal{J}_{\alpha}^j E_n g_j E_m^{\top} - \sum_k E_n \tilde{g}_k E_m^{\top} \mathcal{J}_{\beta}^{\top}. \quad (2.22)$$

Then there exists a polynomial $f \in \mathbb{H}[z]$ satisfying conditions (2.16) if and only if the Sylvester equation

$$\mathcal{J}_{\alpha} X - X \mathcal{J}_{\beta}^{\top} = C \quad (2.23)$$

has a solution $X = [x_{ij}] \in \mathbb{H}^{n \times m}$. If this is the case, all polynomials $f \in \mathbb{H}[z]$ satisfying conditions (2.16) are parametrized by the formula

$$f = g + \sum_{j=1}^m p \cdot x_{n,1} \cdot \rho_{\beta_j} \rho_{\beta_{j-1}} \cdots \rho_{\beta_1} + p \cdot h \cdot \tilde{p} \quad (2.24)$$

when $[x_{n,1} \ \dots \ x_{n,m}]$ is the bottom row of an X satisfying (2.23) and h varies in $\mathbb{H}[z]$. The polynomial f in (2.24) can be written in the form

$$f = \tilde{g} + \sum_{k=1}^n \rho_{\alpha_1} \rho_{\alpha_2} \cdots \rho_{\alpha_k} \cdot x_{k,m} \cdot \tilde{p} + p \cdot h \cdot \tilde{p}. \quad (2.25)$$

The latter theorem is an extension of Theorem 2.2. Formulas (2.19) suggest to conjecture that in more general setting of Theorem 2.3, all solutions f to the interpolation problem should be parametrized by the “whole” solutions X to the equation (2.23) rather than by their bottom rows (in formula (2.24)) or the rightmost columns (in formula (2.25)). However, there is no contradiction here: as we will see in Section 4, any solution X to the equation (2.23) is completely determined by its bottom row or its rightmost column.

Letting $h \equiv 0$ in (2.24) (or in (2.25)) gives all solutions f to the problem (2.16) with $\deg f < \deg p + \deg \tilde{p}$ while letting $g = \tilde{g} \equiv 0$ leads to fairly explicit description of the intersection $\langle p \rangle_{\ell} \cap \langle \tilde{p} \rangle_{\mathbf{r}}$ of two irreducible ideals in $\mathbb{H}[z]$. The case of general polynomials p and \tilde{p} in (2.16) can be reduced to the irreducible case upon making use of the Primary Ideal Decomposition Theorem and gives rise to the Sylvester equation (1.8) with A and B in the block diagonal form with all diagonal blocks of the form (2.21). The formulation of Theorem 2.3 is presented here to explain our interest in the Sylvester equation with A and B of the form (2.21). The proof of the theorem along with reformulations of conditions (2.16) in terms of evaluation functionals is given in [3].

3. QUATERNION SYLVESTER EQUATION: THE REGULAR CASE

Applying the map (2.5) to the equation (1.8) we observe that for any solution $X \in \mathbb{H}^{n \times m}$ of (1.8), the matrix $Y = \varphi(X) \in \mathbb{C}^{2n \times 2m}$ solves the complex Sylvester equation

$$\varphi(A)Y - Y\varphi(B) = \varphi(C). \quad (3.1)$$

By the Sylvester’s criterion and due to equivalence (2.7), the latter equation has a unique solution if and only if $\sigma_{\mathbf{r}}(A) \cap \sigma_{\mathbf{r}}(B) = \emptyset$. If this is the case, we apply the map (2.8) to the equation (3.1). Due to properties (2.9), (2.10) of ψ we get $A\psi(Y) - \psi(Y)B = C$ so that $X = \psi(Y)$ is a solution to (1.8). It is a unique solution since ϕ is injective, so that distinct solutions to (1.8) would have given rise to distinct solutions of (3.1). We thus arrive at the following result (see e.g., [9]):

Theorem 3.1. *Equation (1.8) has a unique solution (for every $C \in \mathbb{H}^{n \times m}$) if and only if $\sigma_{\mathbf{r}}(A) \cap \sigma_{\mathbf{r}}(B) = \emptyset$.*

As we have seen, the unique solution of the equation (1.8) is necessarily of the form $X = \psi(Y)$ where Y is the unique solution to the complex Sylvester equation (3.1). Taking any formula for Y available in literature, one gets a formula for X by letting $X = \psi(Y)$. For example, if we denote by $\mu : \mathbb{C}^{m \times n} \rightarrow \mathbb{C}^{mn \times 1}$ the bijection assigning to the matrix $X = [x_{ij}]$ the the column \mathbf{x} as in (1.2), the formula for the unique solution X of the equation (1.8) suggested by the original Sylvester's approach [17] is

$$X = \psi(\mu^{-1}((\varphi(A) \otimes I_{2m} - I_{2n} \otimes \varphi(B))^{-1} \mu(\phi(C)))). \quad (3.2)$$

The formula with minimal references to complex representations (extending the Jameson's result [10]) has been established in [9, 16]:

$$X = \left(\sum_{k=1}^{2n} \sum_{j=0}^{k-1} a_k A^j C B^{k-j-1} \right) \left(\sum_{j=0}^{2n} a_j B^j \right)^{-1}, \quad (3.3)$$

where $a_0, \dots, a_{2n} \in \mathbb{C}$ are the coefficients of the polynomial

$$\det(\lambda I_{2n} - \phi(A)) = a_0 + a_1 \lambda + \dots + a_{2n} \lambda^{2n}.$$

In the rest of the section we examine how far one can advance making no use of complex representations of quaternion matrices. We start with a very special case where A and B are Jordan blocks and establish the quaternion analog of the formula (1.5). On the other hand, this result generalizes Theorem 1.1.

Theorem 3.2. *Let A and B be Jordan blocks $A = \mathcal{J}_n(\alpha)$ and $B = \mathcal{J}_m^\top(\beta)$ for some $\alpha \not\sim \beta$. Then the equation (1.8) has a unique solution*

$$X = \sum_{k=0}^{n+m-2} \left(\sum_{i=0}^{k+1} (-1)^{k+i} \binom{k+1}{i} \bar{\alpha}^{k-i+1} M_k \beta^i \right) P_{\alpha, \beta}^{-k-1}, \quad (3.4)$$

where $P_{\alpha, \beta}$ is given by (1.9) and

$$M_k = \sum_{\ell=0}^k (-1)^\ell \binom{k}{\ell} F_n^{k-\ell} C (F_m^\top)^\ell \quad \text{for } k = 0, \dots, n+m-2.$$

Proof. Observe that $P_{\alpha, \beta} = \mathcal{X}_{[\alpha]}(\beta)$ (the value of the characteristic polynomial (2.13) at β) and thus, $P_{\alpha, \beta} \neq 0$ if and only if $\alpha \not\sim \beta$. As $A = \alpha I_n + F_n$ and $B = \beta I_m + F_m^\top$ (see (1.3)), the equation (1.8) takes the form

$$\alpha X - X\beta = C - F_n X + X F_m^\top. \quad (3.5)$$

We now subtract the latter equation multiplied by β on the right from the same equation multiplied by $\bar{\alpha}$ on the left:

$$|\alpha|^2 X - (\bar{\alpha} + \alpha) X \beta + X \beta^2 = \bar{\alpha} (C - F_n X + X F_m^\top) - (C - F_n X + X F_m^\top) \beta.$$

Since $|\alpha|^2$ and $(\alpha + \bar{\alpha})$ are real and therefore commute with all quaternions, the latter equality can be written as

$$X P_{\alpha, \beta} = \bar{\alpha} (C - F_n X + X F_m^\top) - (C - F_n X + X F_m^\top) \beta, \quad (3.6)$$

in view of (1.9). Thus, equation (3.6) follows from (3.5). On the other hand, we may subtract the equation (3.6) multiplied by β on the right from the same equation multiplied by α on the left:

$$\alpha X P_{\alpha,\beta} - X P_{\alpha,\beta} \beta = (C - F_n X + X F_m^\top) P_{\alpha,\beta}.$$

Since $P_{\alpha,\beta} \neq 0$ commutes with β , we may cancel $P_{\alpha,\beta}$ in the latter equation arriving at (3.5). Therefore, equations (3.5) and (3.6) are equivalent. Since $P_{\alpha,\beta} \beta = \beta P_{\alpha,\beta}$ and since the matrices F_n and F_m are real, we can iterate the equality (3.6) as follows:

$$\begin{aligned} X P_{\alpha,\beta}^r &= (\bar{\alpha} C - C \beta) P_{\alpha,\beta}^{r-1} - \bar{\alpha} F_n X P_{\alpha,\beta}^{r-1} + \bar{\alpha} X P_{\alpha,\beta}^{r-1} F_m^\top \\ &\quad + F_n X P_{\alpha,\beta}^{r-1} \beta - X P_{\alpha,\beta}^{r-1} F_m^\top \beta \\ &= (\bar{\alpha} C - C \beta) P_{\alpha,\beta}^{r-1} - \bar{\alpha}^2 (F_n C - C F_m^\top) P_{\alpha,\beta}^{r-2} \\ &\quad + 2\bar{\alpha} (F_n C - C F_m^\top) \beta P_{\alpha,\beta}^{r-2} - (F_n C - C F_m^\top) \beta^2 P_{\alpha,\beta}^{r-2} \\ &\quad + F_n (\bar{\alpha}^2 (F_n X - X F_m^\top) - \bar{\alpha} (F_n X - X F_m^\top) \beta) P_{\alpha,\beta}^{r-2} \\ &\quad - (\bar{\alpha}^2 (F_n X - X F_m^\top) - \bar{\alpha} (F_n X - X F_m^\top) \beta) P_{\alpha,\beta}^{r-2} F_m^\top \\ &\quad - F_n (\bar{\alpha} (F_n X - X F_m^\top) \beta - (F_n X - X F_m^\top) \beta^2) P_{\alpha,\beta}^{r-2} \\ &\quad + (\bar{\alpha} (F_n X - X F_m^\top) \beta - (F_n X - X F_m^\top) \beta^2) P_{\alpha,\beta}^{r-2} F_m^\top. \end{aligned}$$

Continuing this iteration and letting $r = m + n - 1$, we get after $r - 1$ steps

$$X P_{\alpha,\beta}^{m+n-1} = \sum_{k=0}^{r-2} \left(\sum_{i=0}^{k+1} (-1)^{k+i} \binom{k+1}{i} \bar{\alpha}^{k-i+1} M_k \beta^i \right) P_{\alpha,\beta}^{m+n-k-2} + R, \quad (3.7)$$

where R contains the terms containing factors F_n^ℓ and $(F_m^\top)^j$ of total degree $\ell + j \geq m + n - 1$. Thus, either $\ell \geq n$ in which case $F_n^\ell = 0$, or $j \geq m$ in which case $(F_m^\top)^j = 0$. Hence all terms in R are zero matrices (i.e., $R = 0$), and formula (3.4) follows immediately from (3.7). Since (3.6) is equivalent to (3.5), each iteration of (3.6) (and in particular, the formula (3.4)) is equivalent to (3.5). Therefore, X of the form (3.4) is a solution to (3.5). The uniqueness of a solution is evident. \square

In Theorems 3.3, 3.4 and 3.8, all restrictions on B will be removed. The next theorem settles the case where $A = \mathcal{J}_\alpha$ is of the form (2.21) with all diagonal entries be in the same conjugacy class. To formulate the theorem, let us note that for the matrices

$$A = \begin{bmatrix} \alpha_1 & 0 & \dots & 0 \\ 1 & \alpha_2 & 0 & \\ 0 & 1 & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & 0 \\ 0 & \dots & 0 & 1 & \alpha_n \end{bmatrix}, \quad A' = \begin{bmatrix} \bar{\alpha}_1 & 0 & \dots & 0 \\ -1 & \bar{\alpha}_2 & 0 & \\ 0 & -1 & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & 0 \\ 0 & \dots & 0 & -1 & \bar{\alpha}_n \end{bmatrix}, \quad (3.8)$$

based on $\alpha_1, \dots, \alpha_n$ from the same conjugacy class, we have

$$A + A' = (\alpha_1 + \bar{\alpha}_1) I_n \quad \text{and} \quad A' A = |\alpha_1|^2 I_n - \tilde{A}, \quad (3.9)$$

where $\tilde{A} = [(\alpha_i - \bar{\alpha}_{i+1})\delta_{i,j+1} + \delta_{i,j+2}]_{i,j=1}^n$; more explicitly:

$$\tilde{A} = \begin{bmatrix} 0 & 0 & \dots & \dots & 0 \\ \alpha_1 - \bar{\alpha}_2 & 0 & & & \\ 1 & \alpha_2 - \bar{\alpha}_3 & \ddots & \ddots & \ddots & \vdots \\ 0 & 1 & \ddots & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & 0 & 0 \\ 0 & \dots & 0 & 1 & \alpha_{n-1} - \bar{\alpha}_n & 0 \end{bmatrix}. \quad (3.10)$$

Relations (3.9) follow from characterization (2.1), due to which $|\alpha_i|^2 = |\alpha_1|^2$ and $\operatorname{Re}\alpha_i = \operatorname{Re}\alpha_1$ for $i = 2, \dots, n$.

Theorem 3.3. *Let $\alpha_1, \dots, \alpha_n$ be the elements from the same conjugacy class $V \subset \mathbb{H}$, let the matrices A , A' and \tilde{A} be defined as in (3.8), (3.10), and let $B \in \mathbb{H}^{m \times m}$ be such that $V \cap \sigma_{\mathbf{r}}(B) = \emptyset$. Then the equation (1.8) has a unique solution*

$$X = \sum_{k=0}^{n-1} \tilde{A}^k (A'C - CB) (\mathcal{X}_V(B))^{-k-1}, \quad (3.11)$$

where \mathcal{X}_V is the characteristic polynomial of the conjugacy class V .

Proof. We multiply equation (1.8) by A' on the left and by B on the right

$$A'AX - A'XB = A'C, \quad AXB - XB^2 = CB,$$

and then subtract the second equality from the first. Due to (3.9), we get

$$|\alpha_1|^2 X - \tilde{A}X - (\alpha_1 + \bar{\alpha}_1)XB + XB^2 = A'C - CB,$$

which can be written in terms of the characteristic polynomial (2.2) as

$$X\mathcal{X}_V(B) = A'C - CB + \tilde{A}X. \quad (3.12)$$

We now iterate (3.12) as in the proof of Theorem 3.2:

$$\begin{aligned} X(\mathcal{X}_V(B))^n &= (A'C - CB)(\mathcal{X}_V(B))^{n-1} + \tilde{A}X(\mathcal{X}_V(B))^{n-1} \\ &= (A'C - CB)(\mathcal{X}_V(B))^{n-1} + \tilde{A}(A'C - CB)(\mathcal{X}_V(B))^{n-2} \\ &\quad + \tilde{A}^2 X(\mathcal{X}_V(B))^{n-2} = \dots \\ &= \sum_{k=0}^{n-1} \tilde{A}^k (A'C - CB) (\mathcal{X}_V(B))^{n-k-1} + \tilde{A}^n X. \end{aligned}$$

The last equality implies (3.11) since $\tilde{A}^n = 0$ (see (3.10)) and $\mathcal{X}_V(B)$ is invertible, since $V \cap \sigma_{\mathbf{r}}(B) = \emptyset$. \square

We next remove the assumption that all diagonal entries in A are in the same conjugacy class. In this case, we get explicit formulas for the rows of a unique solution X . In what follows, we will use the noncommutative product notation

$$\prod_{i=1}^{\widehat{k}} \gamma_i := \gamma_1 \gamma_2 \cdots \gamma_k, \quad \text{and} \quad \prod_{i=1}^{\widehat{k}} \gamma_i := \gamma_k \cdots \gamma_2 \gamma_1 \quad (3.13)$$

Theorem 3.4. *Let $A = \mathcal{J}_\alpha \in \mathbb{H}^{n \times n}$ be defined as in (2.21), let $B \in \mathbb{H}^{m \times m}$ be such that $\sigma_{\mathbf{r}}(B) \cap [\alpha_k] = \emptyset$ for $k = 1, \dots, n$, and let \tilde{C}_j denote the j -th row of the matrix $C \in \mathbb{H}^{n \times m}$. Then the equation (1.8) has a unique solution*

$$X = \text{Col}_{1 \leq k \leq n} \tilde{X}_k, \quad (3.14)$$

with the rows given by

$$\tilde{X}_k = - \sum_{j=1}^k \left(\rho_{\bar{\alpha}_k} \rho_{\bar{\alpha}_{k-1}} \cdots \rho_{\bar{\alpha}_j} \tilde{C}_j \right)^{er} (B) \cdot \prod_{i=j}^{\widehat{k}} (\mathcal{X}_{[\alpha_i]}(B))^{-1}, \quad (3.15)$$

for $k = 1, \dots, n$, where the polynomials $\rho_{\bar{\alpha}_i}$ and $\mathcal{X}_{[\alpha_i]}$ are defined via formulas (2.14) and (2.13), respectively.

The proof is based on the following two observations.

Remark 3.5. *Let f and g be two quaternion matrix polynomials and let B be a square matrix. Then*

$$(fg)^{er}(B) = (f \cdot g^{er}(B))^{er}(B). \quad (3.16)$$

Indeed, since evaluations (2.12) are linear, we have

$$\begin{aligned} (f \cdot g^{er}(B))^{er}(B) &= (f \cdot \sum_j g_j B^j)^{er}(B) \\ &= \sum_k f_k (\sum_j g_j B^j) B^k \\ &= \sum_i (\sum_{j+k=i} f_k g_j) B_i = (fg)^{er}(B). \end{aligned}$$

Remark 3.6. *Let ρ_α be given by (2.14) for a fixed $\alpha \in \mathbb{H}$ and let*

$$(\rho_\alpha D)^{er}(B) = M \quad (3.17)$$

for some $D, M \in \mathbb{H}^{1 \times m}$ and $B \in \mathbb{H}^{m \times m}$ such that $\sigma_{\mathbf{r}}(B) \cap [\alpha] = \emptyset$. Then

$$D = (\rho_{\bar{\alpha}} M)^{er}(B) \cdot (\mathcal{X}_{[\alpha]}(B))^{-1}. \quad (3.18)$$

Indeed, applying (3.16) to $f = \rho_{\bar{\alpha}}$ and $g = \rho_\alpha D$ and taking into account equalities $\rho_{\bar{\alpha}} \rho_\alpha = \mathcal{X}_{[\alpha]}$ and (3.17), we get

$$(\mathcal{X}_{[\alpha]} D)^{er}(B) = (\rho_{\bar{\alpha}} \cdot (\rho_\alpha D)^{er}(B))^{er}(B) = (\rho_{\bar{\alpha}} M)^{er}(B).$$

Since $\mathcal{X}_{[\alpha]}$ is a polynomial with real coefficients, we have

$$(\mathcal{X}_{[\alpha]} D)^{er}(B) = (D \mathcal{X}_{[\alpha]})^{er}(B) = D \mathcal{X}_{[\alpha]}(B),$$

and (3.18) follows from the two latter equalities, once we recall that the matrix $\mathcal{X}_{[\alpha_1]}(B)$ is invertible (as $[\alpha] \cap \sigma_{\mathbf{r}}(B) = \emptyset$).

Proof of Theorem 3.4. Equating the corresponding rows in the equation (1.8) we get, due to (2.21) and (3.14), the system

$$\alpha_1 \tilde{X}_1 - \tilde{X}_1 B = \tilde{C}_1, \quad \alpha_k \tilde{X}_k - \tilde{X}_k B = \tilde{C}_k - \tilde{X}_{k-1} \quad (k = 2, \dots, n),$$

which is equivalent to (1.8). The latter equalities can be written in terms of right evaluations as follows:

$$\left(\rho_{\alpha_1} \tilde{X}_1 \right)^{e_r} (B) = -\tilde{C}_1, \quad \left(\rho_{\alpha_k} \tilde{X}_k \right)^{e_r} (B) = \tilde{X}_{k-1} - \tilde{C}_k \quad (k = 2, \dots, n). \quad (3.19)$$

Making use of Remark 3.6, we solve the leftmost equation in (3.19):

$$\tilde{X}_1 = - \left(\rho_{\alpha_1} \tilde{C}_1 \right)^{e_r} (B) \cdot (\mathcal{X}_{[\alpha_1]}(B))^{-1},$$

confirming formula (3.15) for \tilde{X}_1 . Similarly, we solve the k -th equation in (2.5) for \tilde{X}_k :

$$\tilde{X}_k = \left(\rho_{\alpha_k} (\tilde{X}_{k-1} - \tilde{C}_k) \right)^{e_r} (B) \cdot (\mathcal{X}_{[\alpha_k]}(B))^{-1}. \quad (3.20)$$

Assuming that the formula (3.15) holds for \tilde{X}_{k-1} , we plug it into (3.20):

$$\begin{aligned} \tilde{X}_k &= - \left(\rho_{\alpha_k} \tilde{C}_k \right)^{e_r} (B) \cdot (\mathcal{X}_{[\alpha_k]}(B))^{-1} \\ &\quad - \left(\rho_{\alpha_k} \sum_{j=1}^{k-1} \left(\rho_{\alpha_{k-1}} \cdots \rho_{\alpha_j} \tilde{C}_j \right)^{e_r} (B) \cdot \prod_{i=j}^{\widehat{k-1}} (\mathcal{X}_{[\alpha_i]}(B))^{-1} \right)^{e_r} (B) \\ &\quad \cdot (\mathcal{X}_{[\alpha_k]}(B))^{-1} \\ &= - \left(\rho_{\alpha_k} \tilde{C}_k \right)^{e_r} (B) \cdot (\mathcal{X}_{[\alpha_k]}(B))^{-1} \\ &\quad - \sum_{j=1}^{k-1} \left(\rho_{\alpha_k} \rho_{\alpha_{k-1}} \cdots \rho_{\alpha_j} \tilde{C}_j \right)^{e_r} (B) \cdot \prod_{i=j}^{\widehat{k}} (\mathcal{X}_{[\alpha_i]}(B))^{-1} \\ &= - \sum_{j=1}^k \left(\rho_{\alpha_k} \cdots \rho_{\alpha_j} \tilde{C}_j \right)^{e_r} (B) \cdot \prod_{i=j}^{\widehat{k}} (\mathcal{X}_{[\alpha_i]}(B))^{-1} \end{aligned}$$

and the induction argument completes the proof of formulas (3.15). Note that the second equality in the last calculation followed by Remark 3.5 applied to polynomials $f = \rho_{\alpha_k}$ and $g = \rho_{\alpha_{k-1}} \cdots \rho_{\alpha_j} \tilde{C}_j$ and since the characteristic polynomial $\mathcal{X}_{[\alpha_i]}$ is in $\mathbb{R}[z]$ for $i = 1, \dots, n$. \square

Remark 3.7. If the matrix $A = \mathcal{J}_\alpha$ is based on the elements $\alpha_1, \dots, \alpha_n$ from the same conjugacy class V , then formulas (3.15) simplify to

$$\tilde{X}_k = - \sum_{j=1}^k \left(\rho_{\alpha_k} \rho_{\alpha_{k-1}} \cdots \rho_{\alpha_j} \tilde{C}_j \right)^{e_r} (B) \cdot (\mathcal{X}_V(B))^{j-k-1}.$$

Making use of Remark 3.6, one can get the formula for the unique solution of the equation (1.8) in the case where A is lower triangular.

Theorem 3.8. *Let $A = [\alpha_{i,j}]_{i,j=1}^n$ be a lower triangular matrix ($\alpha_{i,j} = 0$ for $i < j$) and let $B \in \mathbb{H}^{m \times m}$ be such that $\sigma_{\mathbf{r}}(B) \cap [\alpha_{k,k}] = \emptyset$ for $k = 1, \dots, n$. Then the equation (1.8) has a unique solution X of the form (3.14) with the rows defined recursively by*

$$\begin{aligned} \tilde{X}_1 &= - \left(\rho_{\tilde{\alpha}_{1,1}} \tilde{C}_1 \right)^{e_r} (B) \cdot (\mathcal{X}_{[\alpha_{1,1}]}(B))^{-1}, \\ \tilde{X}_k &= - \left(\rho_{\tilde{\alpha}_{k,k}} \tilde{C}_k \right)^{e_r} (B) \cdot (\mathcal{X}_{[\alpha_{k,k}]}(B))^{-1} \\ &\quad + \sum_{j=1}^{k-1} \left(\rho_{\tilde{\alpha}_{k,k}} \alpha_{k,j} \tilde{X}_j \right)^{e_r} (B) \cdot (\mathcal{X}_{[\alpha_{k,k}]}(B))^{-1} \quad \text{for } k = 2, \dots, n. \end{aligned} \quad (3.21)$$

Proof. Equating the corresponding rows in the equation (1.8) we get, due to the lower triangular structure of A , the equations which can be written in terms of right evaluations as

$$\left(\rho_{\alpha_1} \tilde{X}_1 \right)^{e_r} (B) = -\tilde{C}_1, \quad \left(\rho_{\alpha_k} \tilde{X}_k \right)^{e_r} (B) = -\tilde{C}_k + \sum_{j=1}^{k-1} \alpha_{k,j} \tilde{X}_j$$

for $k = 2, \dots, n$. The rest follows by Remark 3.6. \square

To keep the presentation symmetric we conclude with the “column” version of the last theorem.

Theorem 3.9. *Let $B = [\beta_{i,j}]_{i,j=1}^m$ be an upper triangular matrix ($\beta_{i,j} = 0$ for $i > j$) and let $A \in \mathbb{H}^{n \times n}$ be such that $\sigma_{\mathbf{r}}(A) \cap [\beta_{k,k}] = \emptyset$ for $k = 1, \dots, m$ and let $C = [C_1 \ \dots \ C_m]$. Then the equation (1.8) has a unique solution $X = [X_1 \ \dots \ X_m]$ with the columns defined recursively by*

$$\begin{aligned} X_1 &= (\mathcal{X}_{[\beta_{1,1}]}(A))^{-1} \cdot \left(C_1 \rho_{\tilde{\beta}_{1,1}} \right)^{e_\ell} (A), \\ X_k &= (\mathcal{X}_{[\beta_{k,k}]}(A))^{-1} \cdot \left(C_k \bar{\rho}_{\tilde{\beta}_{k,k}} \right)^{e_\ell} (A) \\ &\quad + (\mathcal{X}_{[\beta_{k,k}]}(A))^{-1} \sum_{j=1}^{k-1} \left(X_j \beta_{j,k} \rho_{\tilde{\beta}_{k,k}} \right)^{e_\ell} (A) \quad \text{for } k = 2, \dots, m. \end{aligned} \quad (3.22)$$

Proof. Let us observe that the left and right evaluations (2.12) are related as follows: for any $\alpha \in \mathbb{H}$, $D \in \mathbb{H}^{1 \times m}$ and $B \in \mathbb{H}^{m \times m}$,

$$(\rho_{\alpha} C)^{e_r} (B) = CB - \alpha C = (B^* C^* - C^* \bar{\alpha})^* = [(C^* \rho_{\bar{\alpha}})^{e_\ell} (B^*)]^*. \quad (3.23)$$

Taking adjoints (1.8) we get the equation

$$B^* X^* - X^* A^* = -C^*, \quad (3.24)$$

and since the matrix B^* is lower triangular, we can apply Theorem 3.8 (with A , B , C replaced by B^* , A^* and $-C^*$, respectively) to get recursive formulas for the rows of the matrix X^* . Taking adjoints in these formulas and making use of relations (3.23), we get (3.22). \square

To conclude, we remark that in case $B = \mathcal{J}_\beta^\top$ where \mathcal{J}_β^\top is of the form (2.21), the recursion (3.22) can be solved to produce formulas

$$X_k = \sum_{j=1}^k \prod_{i=j}^{\widehat{k}} (\mathcal{X}_{[\beta_i]}(A))^{-1} \cdot \left(C_j \rho_{\beta_j} \rho_{\beta_{j+1}} \cdots \rho_{\beta_k} \right)^{e_\ell} (A) \quad (1 \leq k \leq m).$$

4. SINGULAR CASE: THE SOLVABILITY CRITERION

In this section, we consider the Sylvester equation (1.8) in the case where $\sigma_{\mathbf{r}}(A) \cap \sigma_{\mathbf{r}}(B) \neq \emptyset$. At certain level, this case also can be handled by making use of complex representation of quaternion matrices: one can pass from (1.8) to the equivalent complex Sylvester equation (3.1), then use any available method (e.g., Sylvester's tensor-product approach recalled at the beginning of the paper) to describe all its solutions, and then claim that the formula $X = \psi(Y)$ describes all solutions to the Sylvester equation (1.8) when Y runs through the set of all solutions to the equation (3.1). This approach is not quite satisfactory, partly because the map ψ (2.8) is not injective. Presumably, some quantitative results still can be obtained on this way: to define the number of independent conditions (scalar equalities) which are necessary and sufficient for the equation (1.8) to have a solution along with the number of independent free parameters needed to parametrize the solution set. In this section we will obtain more definitive results of this sort (explicit solvability conditions and parametrization of all solutions) in the case where the matrices

$$A = \mathcal{J}_\alpha = [\alpha_i \delta_{i,j} + \delta_{i,j+1}]_{i,j=1}^n, \quad B = \mathcal{J}_\beta^\top = [\beta_i \delta_{i,j} + \delta_{i+1,j}]_{i,j=1}^m \quad (4.1)$$

are based on spherical chains $\alpha = (\alpha_1, \dots, \alpha_n)$ and $\beta = (\beta_1, \dots, \beta_m)$. We assume without loss of generality that $n \geq m$ and start with the scalar equation

$$\alpha x - x\beta = c, \quad \text{where } \alpha \sim \beta. \quad (4.2)$$

Let us recall that a unit element $I \in \mathbb{H}$ is purely imaginary if and only if $I^2 = -1$. Therefore, the characterization (2.1) can be reformulated as follows: $\alpha \sim \beta$ if and only if α and β can be written as

$$\alpha = x + yI, \quad \beta = x + y\tilde{I} \quad (x \in \mathbb{R}, y > 0, I^2 = \tilde{I}^2 = -1). \quad (4.3)$$

Since \mathbb{H} is a (four-dimensional) vector space over \mathbb{R} , we may define orthogonal complements with respect to the usual euclidean metric in \mathbb{R}^4 . For α and β as in (4.3), we define the plane (the two-dimensional subspace of $\mathbb{H} \cong \mathbb{R}^4$) $\Pi_{\alpha,\beta}$ via the formula

$$\Pi_{\alpha,\beta} = \begin{cases} \text{span}\{1, I\} = \{u + vI : u, v \in \mathbb{R}\}, & \text{if } \beta = \alpha, \\ (\text{span}\{1, I\})^\perp, & \text{if } \beta = \bar{\alpha}, \\ \text{span}\{I + \tilde{I}, 1 - I\tilde{I}\}, & \text{if } \beta \neq \alpha, \bar{\alpha}. \end{cases} \quad (4.4)$$

Since $\bar{\alpha} = x - yI$, it follows that $\Pi_{\bar{\alpha},\bar{\alpha}} = \Pi_{\alpha,\alpha}$, $\Pi_{\bar{\alpha},\alpha} = \Pi_{\alpha,\bar{\alpha}}$ and

$$\Pi_{\bar{\alpha},\beta} = \text{span}\{I - \tilde{I}, 1 + I\tilde{I}\} \quad \text{if } \beta \neq \alpha, \bar{\alpha}. \quad (4.5)$$

Lemma 4.1 ([1]). *Given $\alpha \sim \beta$, the solution set of the homogeneous Sylvester equation $\alpha x = x\beta$ coincides with $\Pi_{\alpha,\beta}$ given in (4.4). Furthermore, the non-homogeneous*

equation (4.2) has a solution if and only if $\bar{\alpha}c = c\beta$ (equivalently, $c \in \Pi_{\bar{\alpha},\beta}$) in which case the solution set is the affine plane

$$(2\text{Im}(\alpha))^{-1}c + \Pi_{\alpha,\beta} = -c(2\text{Im}(\beta))^{-1} + \Pi_{\alpha,\beta}.$$

We now proceed to the matrix equation

$$AX - XB = C, \quad X = [x_{ij}] = [X_1 \ \dots \ X_m], \quad (4.6)$$

where A and B of the form (4.1) are based on the spherical chains α and β from the same conjugacy class V and where we assume without loss of generality that $m \leq n$. With the given matrices A , B and C , we associate the matrix

$$D := CB - A'C \quad (4.7)$$

where A' is given in (3.8). In more detail, if $C = [c_{i,j}]_{i=1,\dots,n}^{j=1,\dots,m}$ and if we let $c_{i,0} = c_{0,j} = 0$ for all i, j , then

$$D = [d_{i,j}]_{i=1,\dots,n}^{j=1,\dots,m}, \quad d_{i,j} = c_{i,j}\beta_j - \bar{\alpha}_i c_{i,j} + c_{i,j-1} + c_{i-1,j}. \quad (4.8)$$

We next introduce the elements $\Gamma_{k,j} \in \mathbb{H}$ by the following double recursion

$$\Gamma_{k,j} = (\alpha_k - \bar{\alpha}_{k+1})^{-1} [d_{k+1,j} + \Gamma_{k+1,j-2} - \Gamma_{k-1,j} - \Gamma_{k+1,j-1}(\bar{\beta}_j - \beta_{j-1})] \quad (4.9)$$

with the initial conditions

$$\Gamma_{-1,j} = \Gamma_{0,j} = \Gamma_{k,0} = 0 \quad \text{for all } k, j \geq 1. \quad (4.10)$$

It is clear from (4.9) that the assumption that α is a spherical chain (i.e., that $\alpha_k \neq \bar{\alpha}_{k+1}$) is essential. We make several further comments.

Remark 4.2. (1) Recursion (4.9) determines $\Gamma_{k,j}$ for all positive $k < n$ and $j \leq m$ such that $k + j \leq n$.

(2) Any element $\Gamma_{k,1}$ from the first “column” is determined by the elements $\Gamma_{r,1}$ ($1 \leq r < k$).

(3) Any element in the ℓ -th counter-diagonal $\mathcal{D}_\ell = \{\Gamma_{k,j} : k + j = \ell + 1\}$ can be expressed in terms of the elements from the previous counter-diagonal $\mathcal{D}_{\ell-1} = \{\Gamma_{k,j} : k + j = \ell\}$ and one fixed element in \mathcal{D}_ℓ . The latter follows from the formula (4.9) since $\beta_j \neq \bar{\beta}_j$ (β is a spherical chain).

Lemma 4.3. Let $\Gamma_{k,j}$ be defined as in (4.8)–(4.10) and let

$$S_j = d_{1,j} + \Gamma_{1,j-1}(\beta_{j-1} - \bar{\beta}_j) + \Gamma_{1,j-2} \quad (j = 1, \dots, m). \quad (4.11)$$

If $S_j = 0$ for $j = 1, \dots, m$, then

$$\Delta_{k,j} := \alpha_k \Gamma_{k,j} - \Gamma_{k,j} \beta_j - c_{k,j} + \Gamma_{k-1,j} - \Gamma_{k,j-1} = 0 \quad (4.12)$$

for $k < n$ and $j \leq m$ such that $k + j \leq n$.

Proof. We first observe that for $\alpha \sim \beta$,

$$\alpha(\alpha - \bar{\beta})^{-1} = (\alpha - \bar{\beta})^{-1}\beta, \quad \alpha - \bar{\beta} = \beta - \bar{\alpha}. \quad (4.13)$$

By (4.8)–(4.10),

$$\begin{aligned}\Gamma_{k,1} &= (\alpha_k - \bar{\alpha}_{k+1})^{-1} [d_{k+1,1} - \Gamma_{k-1,1}] \\ &= (\alpha_k - \bar{\alpha}_{k+1})^{-1} [c_{k+1,1}\beta_1 - \bar{\alpha}_{k+1}c_{k+1,1} + c_{k,1} - \Gamma_{k-1,1}].\end{aligned}\quad (4.14)$$

Making use of the first equality (with $\alpha = \alpha_k$ and $\beta = \alpha_{k+1}$), we have, on account of (4.7) and (4.14),

$$\begin{aligned}\Delta_{k,1} &= \alpha_k \Gamma_{k,1} - \Gamma_{k,1}\beta_1 - c_{k,1} + \Gamma_{k-1,1} \\ &= (\alpha_k - \bar{\alpha}_{k+1})^{-1} [\alpha_{k+1} (c_{k+1,1}\beta_1 - \bar{\alpha}_{k+1}c_{k+1,1} + c_{k,1} - \Gamma_{k-1,1}) \\ &\quad - (c_{k+1,1}\beta_1 - \bar{\alpha}_{k+1}c_{k+1,1} + c_{k,1} - \Gamma_{k-1,1})\beta_1] - c_{k,1} + \Gamma_{k-1,1} \\ &= (\alpha_k - \bar{\alpha}_{k+1})^{-1} [-c_{k+1}\mathcal{X}_{[\alpha_{k+1}]}(\beta_1) + \bar{\alpha}_k c_{k,1} - c_{k,1}\beta_1 \\ &\quad - \bar{\alpha}_k \Gamma_{k-1,1} + \Gamma_{k-1,1}\beta_1] \\ &= (\bar{\alpha}_{k+1} - \alpha_k)^{-1} [d_{k,1} - c_{k-1,1} + \bar{\alpha}_k \Gamma_{k-1,1} - \Gamma_{k-1,j}\beta_1],\end{aligned}\quad (4.15)$$

where the last equality follows since $\beta_j \sim \alpha_{k+1}$ so that $\mathcal{X}_{[\alpha_{k+1}]}(\beta_j) = 0$, and due to formula (4.8) for $d_{k,1}$. Letting $k = 1$ in (4.15) and taking (4.6) into account, we get

$$\Delta_{1,1} = (\bar{\alpha}_2 - \alpha_1)^{-1} d_{1,1} = (\bar{\alpha}_2 - \alpha_1)^{-1} S_1. \quad (4.16)$$

For $k > 1$, we have

$$\begin{aligned}\bar{\alpha}_k \Gamma_{k-1,j} - \Gamma_{k-1,j}\beta_j &= \Delta_{k-1,1} + (\bar{\alpha}_k - \alpha_{k-1})\Gamma_{k-1,1} + c_{k-1,1} - \Gamma_{k-2,1} \\ &= \Delta_{k-1,1} + c_{k-1,1} - d_{k,1}\end{aligned}$$

where the first equality follows from formula (4.12) for $\Delta_{k-1,1}$, and the second equality follows from (4.14) (with k replaced by $k-1$). Combining the latter equality with (4.15) gives

$$\Delta_{k,1} = (\bar{\alpha}_{k+1} - \alpha_k)^{-1} \Delta_{k-1,1},$$

from which, on account of (4.16), we recursively obtain

$$\Delta_{k,1} = (\bar{\alpha}_{k+1} - \alpha_k)^{-1} (\bar{\alpha}_k - \alpha_{k-1})^{-1} \cdots (\bar{\alpha}_2 - \alpha_1)^{-1} S_1. \quad (4.17)$$

Since $S_1 = 0$, it follows that

$$\Delta_{k,1} := \alpha_k \Gamma_{k,1} - \Gamma_{k,1}\beta_1 - c_{k,1} + \Gamma_{k-1,1} = 0 \quad (k = 1, \dots, n-1). \quad (4.18)$$

We now assume that $j \geq 2$. Making use of the first equality (with $\alpha = \alpha_k$ and $\beta = \alpha_{k+1}$), we have, on account of (4.9),

$$\begin{aligned}\alpha_k \Gamma_{k,j} - \Gamma_{k,j}\beta_j & \\ &= \alpha_k (\alpha_k - \bar{\alpha}_{k+1})^{-1} [d_{k+1,j} + \Gamma_{k+1,j-2} - \Gamma_{k-1,j} - \Gamma_{k+1,j-1}(\bar{\beta}_j - \beta_{j-1})] \\ &\quad - (\alpha_k - \bar{\alpha}_{k+1})^{-1} [d_{k+1,j} + \Gamma_{k+1,j-2} - \Gamma_{k-1,j} - \Gamma_{k+1,j-1}(\bar{\beta}_j - \beta_{j-1})] \beta_j \\ &= (\alpha_k - \bar{\alpha}_{k+1})^{-1} [\alpha_{k+1} d_{k+1,j} - d_{k+1,j}\beta_j + \alpha_{k+1} \Gamma_{k+1,j-2} - \Gamma_{k+1,j-2}\beta_j \\ &\quad - \alpha_{k+1} \Gamma_{k-1,j} + \Gamma_{k-1,j}\beta_j - (\alpha_{k+1} \Gamma_{k+1,j-1} - \Gamma_{k+1,j-1}\beta_{j-1})(\bar{\beta}_j - \beta_{j-1})].\end{aligned}\quad (4.19)$$

Observe that in view of (4.8) and (2.13),

$$\begin{aligned}
\alpha_{k+1}d_{k+1,j} - d_{k+1,j}\beta_j &= \alpha_{k+1}(c_{k+1,j}\beta_j - \bar{\alpha}_{k+1}c_{k+1,j} + c_{k+1,j-1} + c_{k,j}) \\
&\quad - (c_{k+1,j}\beta_j - \bar{\alpha}_{k+1}c_{k+1,j} + c_{k+1,j-1} + c_{k,j})\beta_j \\
&= c_{k+1,j}\mathcal{X}_{[\alpha_{k+1}]}(\beta_j) + \alpha_{k+1}(c_{k+1,j-1} + c_{k,j}) \\
&\quad - (c_{k+1,j-1} + c_{k,j})\beta_j \\
&= \alpha_{k+1}(c_{k+1,j-1} + c_{k,j}) - (c_{k+1,j-1} + c_{k,j})\beta_j,
\end{aligned}$$

where the last equality follows since $\beta_j \sim \alpha_{k+1}$ so that $\mathcal{X}_{[\alpha_{k+1}]}(\beta_j) = 0$. We next observe equalities

$$\begin{aligned}
\alpha_{k+1}c_{k+1,j-1} - c_{k+1,j-1}\bar{\beta}_{j-1} &= c_{k+1,j-1}\beta_{j-1} - \bar{\alpha}_{k+1}c_{k+1,j-1} \\
&= d_{k+1,j-1} - c_{k+1,j-2} - c_{k,j-1} \\
\bar{\alpha}_k c_{k,j} - c_{k,j}\beta_j &= c_{k,j-1} + c_{k-1,j} - d_{k,j},
\end{aligned}$$

which follow from (4.8) and the fact that the elements from the same conjugacy class have the same real part. Combining the three last equalities gives

$$\begin{aligned}
\alpha_{k+1}d_{k+1,j} - d_{k+1,j}\beta_j &= d_{k+1,j-1} - d_{k,j} - c_{k+1,j-2} + c_{k-1,j} \\
&\quad + (\alpha_k - \bar{\alpha}_{k+1})c_{k,j} + c_{k+1,j-1}(\bar{\beta}_j - \beta_{j-1}).
\end{aligned}$$

We now substitute the latter equality into (4.19) and then (4.19) into the definition (4.12) of $\Delta_{k,j}$ to conclude

$$\begin{aligned}
\Delta_{k,j} &= \alpha_k \Gamma_{k,j} - \Gamma_{k,j}\beta_j - c_{k,j} + \Gamma_{k-1,j} - \Gamma_{k,j-1} \\
&= (\alpha_k - \bar{\alpha}_{k+1})^{-1} [d_{k+1,j-1} - d_{k,j} - c_{k+1,j-2} + c_{k-1,j} \\
&\quad + (\alpha_k - \bar{\alpha}_{k+1})c_{k,j} + c_{k+1,j-1}(\bar{\beta}_j - \beta_{j-1}) + \alpha_{k+1}\Gamma_{k+1,j-2} \\
&\quad - \Gamma_{k+1,j-2}\beta_j - \alpha_{k+1}\Gamma_{k-1,j} + \Gamma_{k-1,j}\beta_j \\
&\quad - (\alpha_{k+1}\Gamma_{k+1,j-1} - \Gamma_{k+1,j-1}\beta_{j-1})(\bar{\beta}_j - \beta_{j-1})] \\
&\quad - c_{k,j} + \Gamma_{k-1,j} - \Gamma_{k,j-1} \\
&= (\alpha_k - \bar{\alpha}_{k+1})^{-1} [d_{k+1,j-1} - d_{k,j} - c_{k+1,j-2} + c_{k-1,j} \\
&\quad - \bar{\alpha}_k \Gamma_{k-1,j} + \Gamma_{k-1,j}\beta_j + \alpha_{k+1}\Gamma_{k+1,j-2} - \Gamma_{k+1,j-2}\beta_j \\
&\quad - (\alpha_{k+1}\Gamma_{k+1,j-1} - \Gamma_{k+1,j-1}\beta_{j-1} - c_{k+1,j-1})(\bar{\beta}_j - \beta_{j-1})] - \Gamma_{k,j-1}.
\end{aligned} \tag{4.20}$$

Letting $j = 2$ in the latter equality gives

$$\begin{aligned}
\Delta_{k,2} &= (\alpha_k - \bar{\alpha}_{k+1})^{-1} [d_{k+1,1} - d_{k,2} + c_{k-1,2} - \bar{\alpha}_k \Gamma_{k-1,2} + \Gamma_{k-1,2}\beta_2 \\
&\quad - (\alpha_{k+1}\Gamma_{k+1,1} - \Gamma_{k+1,1}\beta_1 - c_{k+1,1})(\bar{\beta}_2 - \beta_1)] - \Gamma_{k,1}.
\end{aligned} \tag{4.21}$$

Taking into account the first equality in (4.14) and equality

$$\alpha_{k+1}\Gamma_{k+1,1} - \Gamma_{k+1,1}\beta_1 - c_{k+1,1} = -\Gamma_{k,1}$$

which is a consequence of (4.18) (with $k+1$ instead of k), we simplify (4.21) as follows:

$$\begin{aligned}
\Delta_{k,2} &= (\bar{\alpha}_{k+1} - \alpha_k)^{-1} [d_{k,2} - c_{k-1,2} + \bar{\alpha}_k \Gamma_{k-1,2} - \Gamma_{k-1,2}\beta_2 \\
&\quad + \Gamma_{k,1}(\beta_1 - \bar{\beta}_2) - \Gamma_{k-1,1}].
\end{aligned} \tag{4.22}$$

Letting $k = 1$ in (4.22) gives, on account of (4.11),

$$\Delta_{1,2} = (\bar{\alpha}_2 - \alpha_1)^{-1} [d_{1,2} + \Gamma_{1,1}(\beta_1 - \bar{\beta}_2)] = (\bar{\alpha}_2 - \alpha_1)^{-1} S_2. \quad (4.23)$$

On the other hand, if $k \geq 2$, then

$$(\alpha_{k-1} - \bar{\alpha}_k) \Gamma_{k-1,2} = d_{k,2} - \Gamma_{k-2,2} + \Gamma_{k,1}(\beta_1 - \bar{\beta}_2),$$

by formula (4.9) (for k replaced by $k-1$) and therefore, by the definition (4.12) for $\Delta_{k-1,2}$,

$$\begin{aligned} \Delta_{k-1,2} &= (\alpha_{k-1} - \bar{\alpha}_k) \Gamma_{k-1,2} + \bar{\alpha}_k \Gamma_{k-1,2} - \Gamma_{k-1,2} \beta_2 - c_{k-1,2} \\ &\quad + \Gamma_{k-2,2} - \Gamma_{k-1,1} \\ &= d_{k,2} + \Gamma_{k,1}(\beta_1 - \bar{\beta}_2) + \bar{\alpha}_k \Gamma_{k-1,2} - \Gamma_{k-1,2} \beta_2 - c_{k-1,2} - \Gamma_{k-1,1}, \end{aligned}$$

which, being substituted into (4.22), leads us to

$$\Delta_{k,2} = (\bar{\alpha}_{k+1} - \alpha_k)^{-1} \Delta_{k-1,2} \quad \text{for } k = 2, \dots, n-2. \quad (4.24)$$

We now recursively obtain from (4.24) and (4.23) that

$$\Delta_{k,2} = (\bar{\alpha}_{k+1} - \alpha_k)^{-1} (\bar{\alpha}_k - \alpha_{k-1})^{-1} \dots (\bar{\alpha}_2 - \alpha_1)^{-1} S_2,$$

and since $S_2 = 0$, it follows that $\Delta_{k,2} = 0$ for $k = 1, \dots, n-2$.

The rest will be verified by induction in j . Let us assume that

$$\Delta_{k,\ell} = 0 \quad \text{for all } k < n \text{ and } \ell < j \text{ such that } k + \ell \leq n. \quad (4.25)$$

In particular,

$$\begin{aligned} \Delta_{k+1,j-1} &= \alpha_{k+1} \Gamma_{k+1,j-1} - \Gamma_{k+1,j-1} \beta_{j-1} \\ &\quad - c_{k+1,j-1} + \Gamma_{k,j-1} - \Gamma_{k+1,j-2} = 0, \end{aligned}$$

so that

$$\alpha_{k+1} \Gamma_{k+1,j-1} - \Gamma_{k+1,j-1} \beta_{j-1} - c_{k+1,j-1} = \Gamma_{k+1,j-2} - \Gamma_{k,j-1}. \quad (4.26)$$

Substituting the latter equality into (4.20) gives

$$\begin{aligned} \Delta_{k,j} &= (\alpha_k - \bar{\alpha}_{k+1})^{-1} [d_{k+1,j-1} - d_{k,j} - c_{k+1,j-2} + c_{k-1,j} \\ &\quad - \bar{\alpha}_k \Gamma_{k-1,j} + \Gamma_{k-1,j} \beta_j + \alpha_{k+1} \Gamma_{k+1,j-2} - \Gamma_{k+1,j-2} \beta_j \\ &\quad - (\Gamma_{k+1,j-2} - \Gamma_{k,j-1})(\bar{\beta}_j - \beta_{j-1})] - \Gamma_{k,j-1} \\ &= (\alpha_k - \bar{\alpha}_{k+1})^{-1} [d_{k+1,j-1} - d_{k,j} - c_{k+1,j-2} + c_{k-1,j} \\ &\quad - \bar{\alpha}_k \Gamma_{k-1,j} + \Gamma_{k-1,j} \beta_j + \alpha_{k+1} \Gamma_{k+1,j-2} - \Gamma_{k+1,j-2} \bar{\beta}_{j-1} \\ &\quad + \Gamma_{k,j-1}(\bar{\beta}_j - \beta_{j-1}) - (\alpha_k - \bar{\alpha}_{k+1}) \Gamma_{k,j-1}]. \end{aligned} \quad (4.27)$$

By the definition (4.12) of $\Delta_{k+1,j-2}$,

$$\begin{aligned} \alpha_{k+1} \Gamma_{k+1,j-2} - \Gamma_{k+1,j-2} \bar{\beta}_{j-1} &= \Delta_{k+1,j-2} + c_{k+1,j-2} - \Gamma_{k,j-2} \\ &\quad + \Gamma_{k+1,j-3} + \Gamma_{k+1,j-2}(\beta_{j-2} - \bar{\beta}_{j-1}), \end{aligned}$$

and since, by formula (4.9) (with $j-1$ instead of j),

$$\begin{aligned} \Gamma_{k+1,j-2}(\beta_{j-2} - \bar{\beta}_{j-1}) &= (\alpha_k - \bar{\alpha}_{k+1}) \Gamma_{k,j-1} - d_{k+1,j-1} \\ &\quad - \Gamma_{k+1,j-3} + \Gamma_{k-1,j-1}, \end{aligned}$$

combining the two latter equalities with the assumption (4.25) gives

$$\begin{aligned} & \alpha_{k+1}\Gamma_{k+1,j-2} - \Gamma_{k+1,j-2}\bar{\beta}_{j-1} + d_{k+1,j-1} - c_{k+1,j-2} - (\alpha_k - \bar{\alpha}_{k+1})\Gamma_{k,j-1} \\ &= \Gamma_{k-1,j-1} - \Gamma_{k,j-2}. \end{aligned} \quad (4.28)$$

Substituting the latter equality into (4.27) gives

$$\begin{aligned} \Delta_{k,j} &= (\bar{\alpha}_{k+1} - \alpha_k)^{-1} [d_{k,j} - c_{k-1,j} + \bar{\alpha}_k\Gamma_{k-1,j} - \Gamma_{k-1,j}\beta_j \\ &\quad + \Gamma_{k,j-1}(\beta_{j-1} - \bar{\beta}_j) - \Gamma_{k-1,j-1} + \Gamma_{k,j-2}]. \end{aligned} \quad (4.29)$$

Letting $k = 1$ in (4.29) we get, on account of (4.11),

$$\begin{aligned} \Delta_{1,j} &= (\bar{\alpha}_2 - \alpha_1)^{-1} [d_{1,j} + \Gamma_{1,j-1}(\beta_{j-1} - \bar{\beta}_j) + \Gamma_{1,j-2}] \\ &= (\bar{\alpha}_2 - \alpha_1)^{-1} S_j. \end{aligned} \quad (4.30)$$

If $k \geq 2$, then by the definition (4.12) of $\Delta_{k-1,j}$ and by formula (4.9) (with $k-1$ instead of k), we have

$$\begin{aligned} \Delta_{k-1,j} &= (\alpha_{k-1} - \bar{\alpha}_k)\Gamma_{k-1,j} + \bar{\alpha}_k\Gamma_{k-1,j} - \Gamma_{k-1,j}\beta_j \\ &\quad - c_{k-1,j} + \Gamma_{k-2,j} - \Gamma_{k-1,j-1} \\ &= d_{k,j} + \Gamma_{k,j-2} - \Gamma_{k,j-1}(\bar{\beta}_j - \beta_{j-1}) + \bar{\alpha}_k\Gamma_{k-1,j} - \Gamma_{k-1,j}\beta_j \\ &\quad - c_{k-1,j} - \Gamma_{k-1,j-1} \end{aligned} \quad (4.31)$$

which together with (4.30) implies

$$\Delta_{k,j} = (\bar{\alpha}_{k+1} - \alpha_k)^{-1} \Delta_{k-1,j} \quad \text{for } k = 2, \dots, n-j. \quad (4.32)$$

We now recursively obtain from (4.32) and (4.30) that

$$\Delta_{k,j} = (\bar{\alpha}_{k+1} - \alpha_k)^{-1} (\bar{\alpha}_k - \alpha_{k-1})^{-1} \dots (\bar{\alpha}_2 - \alpha_1)^{-1} S_j,$$

and since $S_j = 0$, it follows that $\Delta_{k,j} = 0$ for $k = 1, \dots, n-j$. The induction argument completes the proof of the lemma. \square

Remark 4.4. The elements $\Delta_{k,j}$ defined as in (4.12) satisfy relations

$$\Delta_{k,j} = (\bar{\alpha}_{k+1} - \alpha_k)^{-1} [\Delta_{k+1,j-1}(\bar{\beta}_j - \beta_{j-1}) - \Delta_{k+1,j-2} + \Delta_{k-1,j}] \quad (4.33)$$

for $k \geq 2$ and $j \geq 3$.

Proof. Although we do not assume equalities (4.11) to be in force, formula (4.20) still holds true. Without assumptions (4.25), equality (4.26) gets the extra term $\Delta_{k+1,j-1}$ on the right, so that formula (4.27) takes the form

$$\begin{aligned} \Delta_{k,j} &= (\alpha_k - \bar{\alpha}_{k+1})^{-1} [d_{k+1,j-1} - d_{k,j} - c_{k+1,j-2} + c_{k-1,j} \\ &\quad - \bar{\alpha}_k\Gamma_{k-1,j} + \Gamma_{k-1,j}\beta_j + \alpha_{k+1}\Gamma_{k+1,j-2} - \Gamma_{k+1,j-2}\bar{\beta}_{j-1} \\ &\quad - (\Delta_{k+1,j-1} - \Gamma_{k,j-1})(\bar{\beta}_j - \beta_{j-1}) - (\alpha_k - \bar{\alpha}_{k+1})\Gamma_{k,j-1}]. \end{aligned} \quad (4.34)$$

For $j \geq 3$, we use equality (4.26) with the extra term $\Delta_{k+1,j-2}$ on the right; substituting this modified equality into (4.34) gives the following modification of (4.29):

$$\begin{aligned} \Delta_{k,j} &= (\bar{\alpha}_{k+1} - \alpha_k)^{-1} [d_{k,j} - c_{k-1,j} + \bar{\alpha}_k\Gamma_{k-1,j} - \Gamma_{k-1,j}\beta_j - \Gamma_{k-1,j-1} \\ &\quad + (\Delta_{k+1,j-1} - \Gamma_{k,j-1})(\bar{\beta}_j - \beta_{j-1}) + \Gamma_{k,j-2} - \Delta_{k+1,j-2}]. \end{aligned} \quad (4.35)$$

Combining the latter equality with (4.31) (in case $k \geq 2$) gives (4.33). \square

Since $\Gamma_{1,1} = (\alpha_1 - \bar{\alpha}_2)^{-1}d_{2,1}$, we derive from (4.34)

$$\Delta_{1,2} = (\bar{\alpha}_2 - \alpha_1)^{-1} [d_{1,2} + (\Delta_{2,1} - \Gamma_{1,1})(\bar{\beta}_2 - \beta_1)],$$

which, on account of formulas (4.11) and (4.17) for S_2 and $\Delta_{2,1}$, respectively, can be written as

$$\Delta_{1,2} = (\bar{\alpha}_2 - \alpha_1)^{-1}S_2 + (\bar{\alpha}_2 - \alpha_1)^{-1}(\bar{\alpha}_3 - \alpha_2)^{-1}(\bar{\alpha}_2 - \alpha_1)^{-1}S_1(\bar{\beta}_2 - \beta_1).$$

As expected, the latter formula coincides with (4.23) in case $S_1 = 0$. For $j \geq 3$, we have from (4.35) and formula (4.11) for S_j ,

$$\begin{aligned} \Delta_{1,j} &= (\bar{\alpha}_2 - \alpha_1)^{-1} [d_{1,j} + (\Delta_{2,j-1} - \Gamma_{1,j-1})(\bar{\beta}_j - \beta_{j-1}) + \Gamma_{1,j-2} - \Delta_{2,j-2}] \\ &= (\bar{\alpha}_2 - \alpha_1)^{-1} [S_j + \Delta_{2,j-1}(\bar{\beta}_j - \beta_{j-1}) - \Delta_{2,j-2}], \end{aligned} \quad (4.36)$$

which is the analog of (4.30). It is now clear from (4.33) and (4.36) that $\Delta_{k,j}$ is a two-sided linear combination of S_1, \dots, S_k with left and right coefficients depending respectively, on α and β only. However, explicit formulas for $\Delta_{k,j}$ in terms of S_1, \dots, S_k are quite complicated.

Theorem 4.5. *Given matrices A, B (based on spherical chains $\alpha = (\alpha_1, \dots, \alpha_n)$ and $\beta = (\beta_1, \dots, \beta_m)$ ($m \leq n$) from the same conjugacy class $V \subset \mathbb{H}$) and C as in (4.6), let $d_{i,j}$ and $\Gamma_{i,j}$ be defined as in (4.7)–(4.10) for $k < n$ and $j \leq m$ such that $k + j \leq n$. Then*

(1) *The Sylvester equation (4.6) has a solution if and only if*

$$d_{1,j} + \Gamma_{1,j-1}(\beta_{j-1} - \bar{\beta}_j) + \Gamma_{1,j-2} = 0 \quad \text{for } j = 1, \dots, m. \quad (4.37)$$

(2) *For any solution $X = [x_{i,j}] \in \mathbb{H}^{n \times m}$ to the equation (4.6), $x_{i,j} = \Gamma_{i,j}$ for all $i < n$ and $j \leq m$ such that $i + j \leq n$*

Proof. Let $X \in \mathbb{H}^{n \times m}$ satisfy (4.6). We now verify that

$$\tilde{A}X - X\tilde{B} = D \quad (4.38)$$

where A', \tilde{A} and D are defined in (3.8), (3.10) and (4.7), respectively, and where

$$\tilde{B} = \begin{bmatrix} 0 & \beta_1 - \bar{\beta}_2 & 1 & 0 & \dots & 0 \\ 0 & 0 & \beta_2 - \bar{\beta}_3 & \ddots & \ddots & \vdots \\ & & & & & 0 \\ \vdots & \vdots & \ddots & \ddots & \ddots & 1 \\ & & & & 0 & \beta_{m-1} - \bar{\beta}_m \\ 0 & 0 & \dots & 0 & 0 & 0 \end{bmatrix}. \quad (4.39)$$

We start as in the proof of Theorem 3.3: multiplying the equation (4.6) by A' on the left and by B on the right and then subtracting the second equality from the first we get (3.12). For B of the form (4.1),

$$\begin{aligned} \mathcal{X}_V(B) &= B^2 - 2\text{Re}\beta_1 \cdot B + |\beta_1|^2 I_m \\ &= [\mathcal{X}_V(\beta_i)\delta_{i,j} + (\beta_i - \bar{\beta}_{i+1})\delta_{i+1,j} + \delta_{i+2,j}]_{i,j=1}^m, \end{aligned}$$

and since $\beta_i \in V$ for $i = 1, \dots, n$, we conclude that $\mathcal{X}_V(B) = \tilde{B}$; see (4.39). Now (3.12) takes the form (4.38).

Let X_j and D_j denote the j -th column in X and D , respectively:

$$X_j = \text{Col}_{1 \leq k \leq n} x_{k,j}, \quad D_j = \text{Col}_{1 \leq k \leq n} d_{k,j}.$$

Taking into account the explicit structure (4.39) of \tilde{B} , we now equate the corresponding columns in (4.38):

$$\tilde{A}X_1 = D_1, \tag{4.40}$$

$$\tilde{A}X_2 = D_2 + X_1(\beta_1 - \bar{\beta}_2), \tag{4.41}$$

$$\tilde{A}X_j = D_j + X_{j-1}(\beta_{j-1} - \bar{\beta}_j) + X_{j-2} \quad (3 \leq j \leq m). \tag{4.42}$$

Making use of explicit formula (3.10) for \tilde{A} we equate the top entries in (4.40) and get $d_{1,1} = 0$ (that is, the first condition in (4.37)). Equating other corresponding entries in (4.40) we get

$$\mathbb{A}_n \begin{bmatrix} x_{1,1} \\ x_{2,1} \\ \vdots \\ x_{n-1,1} \end{bmatrix} = \begin{bmatrix} d_{2,1} \\ d_{3,1} \\ \vdots \\ d_{n,1} \end{bmatrix},$$

where

$$\mathbb{A}_n = \begin{bmatrix} \alpha_1 - \bar{\alpha}_2 & 0 & & & 0 \\ 1 & \alpha_2 - \bar{\alpha}_3 & \ddots & \ddots & \vdots \\ 0 & 1 & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & 0 \\ 0 & \dots & 0 & 1 & \alpha_{n-1} - \bar{\alpha}_n \end{bmatrix} \tag{4.43}$$

from which we conclude

$$\begin{aligned} x_{1,1} &= (\alpha_1 - \bar{\alpha}_2)^{-1} d_{2,1}, \\ x_{k,1} &= (\alpha_k - \bar{\alpha}_{k+1})^{-1} (d_{k+1,1} - x_{k-1,1}), \quad 2 \leq k \leq n-1. \end{aligned} \tag{4.44}$$

Comparing (4.44) with (4.14), we see that $x_{1,1} = \Gamma_{1,1}$ and that recursions defining $x_{k,1}$ and $\Gamma_{k,1}$ in terms of $x_{k-1,1}$ and $\Gamma_{k-1,1}$, respectively, are identical. Therefore,

$$x_{k,1} = \Gamma_{k,1} \quad \text{for } k = 1, \dots, n-1. \tag{4.45}$$

Making again use of formula (3.10) for \tilde{A} we equate the top entries in (4.41) and get

$$d_{1,2} + x_{1,1}(\beta_1 - \bar{\beta}_2) = d_{1,2} + \Gamma_{1,1}(\beta_1 - \bar{\beta}_2) = 0$$

(the second equality holds since $x_{1,1} = \Gamma_{1,1}$, by (4.45)), which is the second condition in (4.37). Equating other corresponding entries in (4.40) and taking into account (4.45), we get

$$\mathbb{A}_n \begin{bmatrix} x_{1,2} \\ \vdots \\ x_{n-2,2} \\ x_{n-1,2} \end{bmatrix} = \begin{bmatrix} d_{2,1} \\ \vdots \\ d_{n-1,1} \\ d_{n,1} \end{bmatrix} + \begin{bmatrix} x_{2,1} \\ \vdots \\ x_{n-1,1} \\ x_{n,1} \end{bmatrix} (\beta_1 - \bar{\beta}_2) = \begin{bmatrix} d_{2,1} \\ \vdots \\ d_{n-1,1} \\ d_{n,1} \end{bmatrix} + \begin{bmatrix} \Gamma_{2,1} \\ \vdots \\ \Gamma_{n-1,1} \\ x_{n,1} \end{bmatrix} (\beta_1 - \bar{\beta}_2),$$

where \mathbb{A}_n is given by (4.43). Taking into account the two-diagonal structure (4.43) of \mathbb{A}_n , we derive from the last equation

$$\begin{aligned} x_{1,2} &= (\alpha_1 - \bar{\alpha}_2)^{-1}(d_{2,1} + \Gamma_{1,1}(\beta_1 - \bar{\beta}_2)), \\ x_{k,2} &= (\alpha_k - \bar{\alpha}_{k+1})^{-1}(d_{k+1,1} + \Gamma_{k+1,1}(\beta_1 - \bar{\beta}_2) - \Gamma_{k-1,2}), \quad 2 \leq k \leq n-2. \end{aligned} \quad (4.46)$$

Comparing (4.46) with (4.9), we see that $x_{1,2} = \Gamma_{1,2}$ and that recursions defining $x_{k,2}$ and $\Gamma_{k,2}$ in terms of $x_{k-1,2}$ and $\Gamma_{k-1,2}$, respectively, are identical. Therefore,

$$x_{k,2} = \Gamma_{k,2} \quad \text{for } k = 1, \dots, n-2. \quad (4.47)$$

We now choose an integer ℓ ($3 \leq \ell < m$) and assume that conditions (4.37) hold for all $j = 1, \dots, \ell-1$ and that $x_{k,j} = \Gamma_{k,j}$ for all $j < \ell$ and $k = 1, \dots, n-j$. Equating the top entries in (4.42) (for $j = \ell$) we get

$$0 = d_{1,\ell} + x_{1,\ell-1}(\beta_{\ell-1} - \bar{\beta}_\ell) + x_{1,\ell-2} = d_{1,\ell} + \Gamma_{1,\ell-1}(\beta_{\ell-1} - \bar{\beta}_\ell) + \Gamma_{1,\ell-2}$$

which is the ℓ -th condition in (4.37). Equating all other entries in the same equation gives

$$\mathbb{A}_n \begin{bmatrix} x_{1,\ell} \\ \vdots \\ x_{n-1,\ell} \end{bmatrix} = \begin{bmatrix} d_{2,\ell} \\ \vdots \\ d_{n,\ell} \end{bmatrix} + \begin{bmatrix} x_{2,\ell-1} \\ \vdots \\ x_{n,\ell-1} \end{bmatrix} (\beta_{\ell-1} - \bar{\beta}_\ell) + \begin{bmatrix} x_{2,\ell-2} \\ \vdots \\ x_{n,\ell-2} \end{bmatrix}$$

which implies (similarly to (4.44) and (4.46),

$$\begin{aligned} x_{1,\ell} &= (\alpha_\ell - \bar{\alpha}_{\ell+1})^{-1}(d_{2,\ell} + x_{2,\ell-1}(\beta_1 - \bar{\beta}_2) + x_{2,\ell-2}), \\ x_{k,\ell} &= (\alpha_k - \bar{\alpha}_{k+1})^{-1}(d_{k+1,\ell} + x_{k+1,\ell-1}(\beta_{\ell-1} - \bar{\beta}_\ell) + x_{k+1,\ell-2} - x_{k-1,\ell}) \end{aligned}$$

for $k = 2, \dots, n$. At least for $k \leq n - \ell$, we may use assumptions $x_{k,j} = \Gamma_{k,j}$ to rewrite the above equalities (more precisely, the right hand side expressions in the above equalities) as

$$\begin{aligned} x_{1,\ell} &= (\alpha_\ell - \bar{\alpha}_{\ell+1})^{-1}(d_{2,\ell} + \Gamma_{2,\ell-1}(\beta_1 - \bar{\beta}_2) + \Gamma_{2,\ell-2}), \\ x_{k,\ell} &= (\alpha_k - \bar{\alpha}_{k+1})^{-1}(d_{k+1,\ell} + \Gamma_{k+1,1}(\beta_{\ell-1} - \bar{\beta}_\ell) + \Gamma_{k+1,\ell-2} - x_{k-1,\ell}). \end{aligned} \quad (4.48)$$

We again compare (4.47) with (4.9) to see that $x_{1,\ell} = \Gamma_{1,\ell}$ and that $x_{k,\ell}$ and $\Gamma_{k,\ell}$ are defined by the same recursion. Therefore $x_{k,\ell} = \Gamma_{k,\ell}$ for $k = 1, \dots, n - \ell$, and then we conclude by induction that equalities (4.37) hold for $j = 1, \dots, m$ and that $x_{k,j} = \Gamma_{k,j}$ for all $k < n$ and $j < m$ such that $k + j \leq n$.

It remains to show that conditions (4.37) are sufficient for the equation (4.6) to have a solution. Let us extend the spherical chain $\alpha = (\alpha_1, \dots, \alpha_n)$ to a spherical chain $\hat{\alpha} = (\alpha_1, \dots, \alpha_n, \alpha_{n+1}, \dots, \alpha_{n+m})$. We may let, for example, $\alpha_{n+i} = \alpha_n$ for $i = 1, \dots, m$. We then let

$$\hat{A} = \mathcal{J}_{\hat{\alpha}} = [\alpha_i \delta_{i,j} + \delta_{i,j+1}]_{i,j=1}^{n+m} = \begin{bmatrix} A & 0 \\ * & * \end{bmatrix}$$

and

$$\hat{A}' = [\bar{\alpha}_i \delta_{i,j} - \delta_{i,j+1}]_{i,j=1}^{n+m} = \begin{bmatrix} A' & 0 \\ * & * \end{bmatrix}$$

to be the corresponding extensions of A and A' (see (3.8)) and we arbitrarily extend the given matrix C to

$$\widehat{C} = \begin{bmatrix} C \\ C' \end{bmatrix} = [c_{i,j}]_{i=1,\dots,n+m}^{j=1,\dots,m} \in \mathbb{H}^{(n+m) \times m}$$

(we may choose $C' = 0$). We then consider the matrix $\widehat{D}' = \widehat{C}B - \widehat{A}'\widehat{C} \in \mathbb{H}^{(n+m) \times n}$ which is indeed an extension of the matrix (4.7):

$$\begin{aligned} \widehat{D}' &= \widehat{C}B - \widehat{A}'\widehat{C} = \begin{bmatrix} C \\ C' \end{bmatrix} B - \begin{bmatrix} A' & 0 \\ * & * \end{bmatrix} \begin{bmatrix} C \\ C' \end{bmatrix} \\ &= \begin{bmatrix} CB - A'C \\ D' \end{bmatrix} = \begin{bmatrix} D \\ D' \end{bmatrix} = [d_{i,j}]_{i=1,\dots,n+m}^{j=1,\dots,m}. \end{aligned}$$

We now use recursions (4.9), (4.10) to define the elements $\Gamma_{k,j}$ for all positive $k < n + m$ and $j \leq m$ such that $k + j \leq n + m$; see Remark 4.2 (part (1)). By Remark 4.2 (part (3)) and by formula (4.9), the elements $\Gamma_{k,j}$ (for $k + j \leq n$) are completely determined by the entries of the original matrices A , B and C . Since equalities (4.37) hold (i.e., the elements S_j in (4.11) are equal to zero for $j = 1, \dots, m$, Lemma 4.3 applies to the extended set $\{\Gamma_{k,j} : k + j \leq n + m\}$ and guarantees that equalities (4.12) hold for all $k < n + m$ and $j \leq m$ such that $k + j \leq n + m$. In particular, equalities (4.12) hold for all $1 \leq k \leq n$ and $1 \leq j \leq m$. Furthermore, if we let

$$\Gamma = [\Gamma_{k,j}]_{k=1,\dots,n}^{j=1,\dots,m} \in \mathbb{H}^{n \times m}, \quad (4.49)$$

then (as it is readily seen from (4.1)) $\Delta_{k,j}$ defined as in (4.12) is the (k, j) -entry in the matrix

$$\Delta = [\Delta_{k,j}]_{k,j} = A\Gamma - \Gamma B - C.$$

Equalities (4.12) for all $1 \leq k \leq n$ and $1 \leq j \leq m$ mean that $\Delta = 0$, i.e., that Γ is a solution to the equation (4.6). \square

The last theorem suggested an algorithm for constructing a particular solution for a solvable singular Sylvester equation. We record it for the convenience of future reference.

Algorithm 4.6. Given A and B of the form (4.1) where $\alpha = (\alpha_1, \dots, \alpha_n)$ and $\beta = (\beta_1, \dots, \beta_n)$ ($m \leq n$) are spherical chains from the same conjugacy class, and given $C = [c_{i,k}] \in \mathbb{H}^{n \times m}$,

- (1) Compute $d_{i,j}$ for $i + j \leq n$ by formula (4.8).
- (2) Compute $\Gamma_{i,j}$ for $i + j \leq n$ by formulas (4.9), (4.10).
- (3) Verify equalities (4.11) for $j = 1, \dots, m$. If they hold true, equation (4.6) has a solution.
- (4) Let $\alpha_{n+i} = \alpha_n$ and $c_{n+i,j} = 0$ for $i, j = 1, \dots, m$.
- (5) Compute $d_{i,j}$ for $n + 1 \leq i + j \leq n + m$ by formula (4.8).
- (6) Compute $\Gamma_{i,j}$ for $n + 1 \leq i + j \leq n + m$ by formula (4.9).

The matrix (4.49) is a particular solution of the Sylvester equation (4.6)

It is of particular interest to write necessary and sufficient conditions (4.37) for solvability of a singular Sylvester equation (4.6) in terms of given matrices A , B and C . We were able to establish such formulas only in the case where A and B

are Jordan blocks (i.e., A and B are based on spherical chains of the form (2.3)). Theorem 4.7 below is the quaternion analog of a result of Ma [13]; see (1.7).

Theorem 4.7. *Let A and B be of the form (4.1) with $\alpha_i = \alpha \sim \beta = \beta_j$ ($1 \leq i \leq n$, $1 \leq j \leq m \leq n$). Given a matrix $C = [c_{k,j}] \in \mathbb{H}^{n \times m}$, let*

$$\begin{aligned} d_{1,1} &= c_{1,1}\beta - \bar{\alpha}c_{1,1}, \\ d_{1,j} &= c_{1,j}\beta - \bar{\alpha}c_{1,j} + c_{1,j-1} \quad (j \geq 2), \\ d_{k,1} &= c_{k,1}\beta - \bar{\alpha}c_{k,1} + c_{k-1,1} \quad (k \geq 2), \\ d_{k,j} &= c_{k,j}\beta - \bar{\alpha}c_{k,j} + c_{k-1,j} + c_{k,j-1} \quad (k, j \geq 2). \end{aligned}$$

Sylvester equation (4.6) has a solution if and only (c.f. with (1.7))

$$\sum_{2\ell \leq j} \text{Im}(\alpha) \cdot d_{2\ell+1,j-2\ell} + \sum_{2\ell \leq j} d_{2\ell,j-2\ell+1} \cdot \text{Im}(\beta) = 0 \quad \text{for } j = 1, \dots, m. \quad (4.50)$$

Proof. We will show that in the present setting, conditions (4.37) are equivalent to (4.50). We first observe that $|\text{Im}(\alpha)| = |\text{Im}(\beta)|$ by characterization (2.1). Therefore, $(\alpha - \bar{\alpha})^2 = (\beta - \bar{\beta})^2$ is a negative number and therefore

$$(\alpha - \bar{\alpha})^{-2}d(\beta - \bar{\beta})^2 = d \quad \text{for any } d \in \mathbb{H}. \quad (4.51)$$

Making use of notation (4.11) and formula (4.9) for $\Gamma_{1,j-1}$ (specified to the present particular case) we get

$$\begin{aligned} (\alpha - \bar{\alpha})S_j &= (\alpha - \bar{\alpha}) [d_{1,j} + \Gamma_{1,j-1}(\beta - \bar{\beta}) + \Gamma_{1,j-2}] \\ &= (\alpha - \bar{\alpha})(d_{1,j} + \Gamma_{1,j-2}) + (d_{2,j-1} + \Gamma_{2,j-3} + \Gamma_{2,j-2}(\beta - \bar{\beta}))(\beta - \bar{\beta}) \\ &= (\alpha - \bar{\alpha})d_{1,j} + d_{2,j-1}(\beta - \bar{\beta}) \\ &\quad + \Gamma_{2,j-3}(\beta - \bar{\beta}) + (\alpha - \bar{\alpha})^2\Gamma_{2,j-2} + (\alpha - \bar{\alpha})\Gamma_{1,j-2}, \end{aligned} \quad (4.52)$$

where the last equality holds due to (4.51). Making subsequent use of formula (4.9) for $\Gamma_{2,j-2}$ and then for $\Gamma_{3,j-3}$, we conclude that the sum of the three rightmost terms in (4.52) equals

$$\begin{aligned} &\Gamma_{2,j-3}(\beta - \bar{\beta}) + (\alpha - \bar{\alpha}) [d_{3,j-2} + \Gamma_{3,j-4} + \Gamma_{3,j-3}(\beta - \bar{\beta})] \\ &= (\alpha - \bar{\alpha}) [d_{3,j-2} + \Gamma_{3,j-4}] + [d_{4,j-3} + \Gamma_{4,j-5} + \Gamma_{4,j-4}(\beta - \bar{\beta})] (\beta - \bar{\beta}) \\ &= (\alpha - \bar{\alpha})d_{3,j-2} + d_{4,j-3}(\beta - \bar{\beta}) \\ &\quad + \Gamma_{4,j-5}(\beta - \bar{\beta}) + (\alpha - \bar{\alpha})^2\Gamma_{4,j-4} + (\alpha - \bar{\alpha})\Gamma_{3,j-4}. \end{aligned}$$

which being substituted into (4.52) gives

$$\begin{aligned} (\alpha - \bar{\alpha})S_j &= (\alpha - \bar{\alpha})(d_{1,j} + d_{3,j-2}) + (d_{2,j-1} + d_{4,j-3})(\beta - \bar{\beta}) \\ &\quad + \Gamma_{4,j-5}(\beta - \bar{\beta}) + (\alpha - \bar{\alpha})^2\Gamma_{4,j-4} + (\alpha - \bar{\alpha})\Gamma_{3,j-4}. \end{aligned} \quad (4.53)$$

The step which led us from (4.52) to (4.53) can be repeated indefinitely. After ℓ iterations we get

$$\begin{aligned}
 (\alpha - \bar{\alpha})S_j = & (\alpha - \bar{\alpha}) \sum_{k=0}^{\ell-1} d_{2k+1,j-2k} + \sum_{k=1}^{\ell} d_{2k,j-2k+1}(\beta - \bar{\beta}) \\
 & + \Gamma_{2\ell,j-2\ell-1}(\beta - \bar{\beta}) + (\alpha - \bar{\alpha})^2 \Gamma_{2\ell,j-2\ell} + (\alpha - \bar{\alpha}) \Gamma_{2\ell-1,j-2\ell}.
 \end{aligned} \tag{4.54}$$

Recall that $\Gamma_{k,j}$ and $d_{k,j}$ are equal zero whenever one of the indices is nonpositive. If we choose ℓ be greater than $j/2$ in (4.53) and take into account only nonzero terms, we conclude that $(\alpha - \bar{\alpha})S_j$ equals the expression on the left side of (4.50). Thus, conditions (4.50) are equivalent (since $\alpha \neq \bar{\alpha}$ to those in (4.37) and applying Theorem 4.5 completes the proof. \square

5. SINGULAR CASE: PARAMETRIZATION OF ALL SOLUTIONS

Since the equation (4.6) is linear, all its solutions X are described by $X = \Gamma + Y$, where Y is the general solution of the homogeneous equation

$$AY - YB = 0. \tag{5.1}$$

we will use notation $\mathcal{D}_\ell(Y) = \{Y_{k,j} : k + j = \ell + 1\}$ for the ℓ -th counter-diagonal of the matrix $Y = [Y_{k,j}]$.

In this final part of the paper, we will present a parametrization of the solution set of the homogeneous equation (5.1) with free parameters $\mu_j \in \Pi_{\alpha_n, \beta_j}$ ($j = 1, \dots, m$) from the planes $\Pi_{\alpha_n, \beta_j} \subset \mathbb{H}$ defined as in (4.4). The latter memberships mean, by Lemma 4.1, that

$$\alpha_n \mu_j = \mu_j \beta_j \quad \text{for } j = 1, \dots, m. \tag{5.2}$$

For any fixed tuple $\boldsymbol{\mu} = (\mu_1, \dots, \mu_m)$ of such parameters, we recursively define the entries $Y_{k,j}$ of the matrix $Y = [Y_{k,j}] \in \mathbb{H}^{n \times m}$ by the formulas

$$Y_{k,j} = 0 \quad \text{for all } k + j \leq n, \quad Y_{n,1} = \mu_1, \tag{5.3}$$

$$\begin{aligned}
 Y_{k,j} = & (\bar{\alpha}_{k+1} - \alpha_k)^{-1} [Y_{k+1,j-1}(\bar{\beta}_j - \beta_{j-1}) - Y_{k+1,j-2} + Y_{k-1,j}] \\
 & (n - m + 1 \leq k \leq n - 1; \quad n - k + 1 \leq j \leq n + m - k),
 \end{aligned} \tag{5.4}$$

$$Y_{n,j+1} = \mu_{j+1} - \sum_{\ell=1}^j \left(\prod_{i=1}^{\ell} (\bar{\alpha}_{n-i+1} - \alpha_{n-i})^{-1} \right) Y_{n-\ell+1,j} \quad (2 \leq j \leq m). \tag{5.5}$$

For the sake of clarity, we display formula (5.5) in a less compact form

$$\begin{aligned}
 Y_{n,j+1} = & \mu_{j+1} - (\bar{\alpha}_n - \alpha_{n-1})^{-1} Y_{n,j} \\
 & - (\bar{\alpha}_n - \alpha_{n-1})^{-1} (\bar{\alpha}_{n-1} - \alpha_{n-2})^{-1} Y_{n-1,j} - \dots \\
 & - (\bar{\alpha}_n - \alpha_{n-1})^{-1} \dots (\bar{\alpha}_{n-j+1} - \alpha_{n-j})^{-1} Y_{n-j+1,j}.
 \end{aligned}$$

The construction of Y goes as follows: with (5.3) in hands, we use formula (5.4) to compute all entries in the counter-diagonal $\mathcal{D}_{n+1}(Y)$, it is not hard to see that

$$Y_{n-j,j+1} = \left(\prod_{i=1}^{\widehat{j}} (\bar{\alpha}_{n-i+1} - \alpha_{n-i})^{-1} \right) \mu_1 \left(\prod_{i=1}^{\widehat{j}} (\bar{\beta}_{i+1} - \beta_i) \right). \quad (5.6)$$

In particular, for $j = 2$ we have

$$Y_{n-1,2} = (\bar{\alpha}_n - \alpha_{n-1})^{-1} \mu_1 (\bar{\beta}_2 - \beta_1) \quad (5.7)$$

and then we use (5.5) to compute

$$Y_{n,2} = \mu_2 - (\bar{\alpha}_n - \alpha_{n-1})^{-1} Y_{n,1} = \mu_2 - (\bar{\alpha}_n - \alpha_{n-1})^{-1} \mu_1, \quad (5.8)$$

and then again (5.4) to compute the entries in $\mathcal{D}_{n+2}(Y)$. Once all counter-diagonals $\mathcal{D}_\ell(Y)$ are found for $r = 2, \dots, n+j+1$ (and in particular, $Y_{k,j}$ are specified for all $k = 1, \dots, n$), we use (5.5) to find $Y_{n,j+1}$ and then (5.4) to compute the next counter-diagonals $\mathcal{D}_{n+j+2}(Y)$.

Theorem 5.1. *Let $\mu = (\mu_1, \dots, \mu_m)$ satisfy (5.2). Then the matrix $Y \in \mathbb{H}^{n \times m}$ constructed as in (5.3)–(5.5) is a solution to the equation (5.1). Moreover, any solution Y to the equation (5.1) arises in this way.*

Proof. The recursion formula (5.4) is the homogeneous version of (4.9) (with all $c_{k,j}$ or, equivalently, all $d_{k,j}$ equal zero). Hence, some conclusions will be obtained by applying the corresponding counter-parts from the previous section. In analogy to the non-homogeneous case, we introduce

$$\Delta_{k,j} := \alpha_k Y_{k,j} - Y_{k,j} \beta_j + Y_{k-1,j} - Y_{k,j-1} \quad (5.9)$$

for $k = 1, \dots, n$ and $j = 1, \dots, m$, and recall relations (4.33) holding for all $k \geq 3$ and $j \geq 2$. We write them equivalently as

$$\Delta_{k-1,j} = (\bar{\alpha}_{k+1} - \alpha_k) \Delta_{k,j} - \Delta_{k+1,j-1} (\bar{\beta}_j - \beta_{j-1}) + \Delta_{k+1,j-2} \quad (5.10)$$

and observe that $\Delta_{k+1,j-2}$ and $\Delta_{k-1,j}$ are the consecutive entries on the $(k+j+1)$ -th counter-diagonal of the matrix $\Delta = [\Delta_{k,j}]$ whereas $\Delta_{k+1,j-1}$ and $\Delta_{k+1,j-2}$ are the entries from the “previous” counter-diagonal $\mathcal{D}_{k+j}(\Delta)$. Thus, relations (5.10) guarantee that once all entries in $\mathcal{D}_r(\Delta)$ are zeros for some $r > n$ and the bottom entry of the next counter-diagonal $\mathcal{D}_{r+1}(\Delta)$ is zero, then all entries in $\mathcal{D}_{r+1}(\Delta)$ are zeros. In other words, the following implication holds true: *for any fixed r ($n < r < n+m$),*

$$\begin{cases} \Delta_{k,r-k} = 0 \ (r-m \leq k \leq n), \\ \Delta_{n,r+1-n} = 0 \end{cases} \implies \Delta_{k,r+1-k} = 0 \ (r-m+1 \leq k \leq n). \quad (5.11)$$

Furthermore, due to (5.3), $\Delta_{k,j} = 0$ for all (k,j) such that $k+j \leq n$. We next observe that due to (5.2),

$$\Delta_{n,1} = \alpha_n \mu_1 - \mu_1 \beta_1 = 0. \quad (5.12)$$

Making use of (5.7), (4.13) and (5.12) we get

$$\begin{aligned}\Delta_{n-1,2} &= \alpha_{n-1}Y_{n-1,2} - Y_{n-1,2}\beta_2 \\ &= \alpha_{n-1}(\bar{\alpha}_n - \alpha_{n-1})^{-1}\mu_1(\bar{\beta}_2 - \beta_1) - (\bar{\alpha}_n - \alpha_{n-1})^{-1}\mu_1(\bar{\beta}_2 - \beta_1)\beta_2 \\ &= (\bar{\alpha}_n - \alpha_{n-1})^{-1}(\alpha_n\mu_1 - \mu_1\beta_1)(\bar{\beta}_2 - \beta_1) = 0.\end{aligned}$$

Since $\Delta_{k,j} = 0$ for all k, j such that $k + j = n$ and since $\Delta_{n-1,2} = 0$, we recursively get from (5.10) that $\Delta_{k,j} = 0$ for all k, j such that $k + j = n + 1$ (i.e., that all the entries in $\mathcal{D}_n(\Delta)$ are zeros). To use induction principle, let us assume that $\Delta_{k,i} = 0$ for all k, i such that $k + i \leq r$ for some $r \geq n + 2$ and show that $\Delta_{k,i} = 0$ also for all k, i such that $k + i = r + 1$. Let $j := n - r$. By the assumption, we have in particular, $\Delta_{k,j} = 0$ for all $k = 1, \dots, n$ which implies, due to formula (5.9), equalities

$$\alpha_k Y_{k,j} - Y_{k,j} \beta_j = Y_{k,j-1} - Y_{k-1,j} \quad \text{for } k = 2, \dots, n. \quad (5.13)$$

Since $\text{Re}(\alpha_k) = \text{Re}(\beta_j)$ (by characterization (2.1)), the latter equalities can be written equivalently as

$$\bar{\alpha}_k Y_{k,j} - Y_{k,j} \bar{\beta}_j = Y_{k-1,j} - Y_{k,j-1} \quad \text{for } k = 2, \dots, n. \quad (5.14)$$

We next compute, for a fixed k ($2 \leq k \leq n - 1$),

$$\begin{aligned}Y_{k,j+1} - Y_{k+1,j} &= (\bar{\alpha}_{k+1} - \alpha_k)^{-1} [Y_{k+1,j}(\bar{\beta}_j - \beta_{j+1}) - Y_{k+1,j-1} + Y_{k-1,j+1} \\ &\quad - (\bar{\alpha}_{k+1} - \alpha_k)Y_{k+1,j}] \\ &= (\bar{\alpha}_{k+1} - \alpha_k)^{-1} [\alpha_k Y_{k+1,j} - Y_{k+1,j} \beta_{j+1} - Y_{k,j} + Y_{k-1,j+1}] \\ &= \alpha_{k+1}(\bar{\alpha}_{k+1} - \alpha_k)^{-1} Y_{k+1,j} - (\bar{\alpha}_{k+1} - \alpha_k)^{-1} Y_{k+1,j} \beta_{j+1} \\ &\quad + (\bar{\alpha}_{k+1} - \alpha_k)^{-1} [Y_{k-1,j+1} - Y_{k,j}],\end{aligned} \quad (5.15)$$

where for the first equality we used formula (5.4) for $Y_{k,j+1}$, the second equality holds due to (5.14) (with k replaced by $(k + 1)$), and the last equality relies on (4.13) (for $\alpha = \alpha_{k+1}$ and $\beta = \alpha_k$). Writing formula (5.14) for $k = n - 1$ and then iterating it once and making use of (4.13), we get

$$\begin{aligned}Y_{n-1,j+1} - Y_{n,j} &= \alpha_n(\bar{\alpha}_n - \alpha_{n-1})^{-1}Y_{n,j} - (\bar{\alpha}_n - \alpha_{n-1})^{-1}Y_{n,j}\beta_{j+1} \\ &\quad + (\bar{\alpha}_n - \alpha_{n-1})^{-1}[\alpha_{n-1}(\bar{\alpha}_{n-1} - \alpha_{n-2})^{-1}Y_{n-1,j} \\ &\quad - (\bar{\alpha}_{n-1} - \alpha_{n-2})^{-1}Y_{n-1,j}\beta_{j+1} \\ &\quad + (\bar{\alpha}_{n-1} - \alpha_{n-2})^{-1}[Y_{n-3,j+1} - Y_{n-2,j}]] \\ &= \alpha_n(\bar{\alpha}_n - \alpha_{n-1})^{-1}Y_{n,j} - (\bar{\alpha}_n - \alpha_{n-1})^{-1}Y_{n,j}\beta_{j+1} \\ &\quad + \alpha_n(\bar{\alpha}_n - \alpha_{n-1})^{-1}(\bar{\alpha}_{n-1} - \alpha_{n-2})^{-1}Y_{n-1,j} \\ &\quad - (\bar{\alpha}_n - \alpha_{n-1})^{-1}(\bar{\alpha}_{n-1} - \alpha_{n-2})^{-1}Y_{n-1,j}\beta_{j+1} \\ &\quad + (\bar{\alpha}_n - \alpha_{n-1})^{-1}(\bar{\alpha}_{n-1} - \alpha_{n-2})^{-1}[Y_{n-3,j+1} - Y_{n-2,j}].\end{aligned}$$

We can continue iterating by invoking the formula (5.14) for $k = n - 3$. After j iterations, we will get to the difference $[Y_{n-j-1,j+1} - Y_{n-j,j}]$ of two entries from the

counter-diagonal $\mathcal{D}_{n+1}(Y)$ which are zeros, by (5.3). Thus, j iterations of formula (5.14) (for $k = n - 1, \dots, n - j$) results in

$$Y_{n-1,j+1} - Y_{n,j} = \alpha_n \left(\sum_{\ell=1}^j \left(\prod_{i=1}^{\ell} (\bar{\alpha}_{n-i+1} - \alpha_{n-i})^{-1} \right) Y_{n-\ell+1,j} \right) - \left(\sum_{\ell=1}^j \left(\prod_{i=1}^{\ell} (\bar{\alpha}_{n-i+1} - \alpha_{n-i})^{-1} \right) Y_{n-\ell+1,j} \right) \beta_{j+1}. \quad (5.16)$$

We now combine latter formula with formulas (5.9), (5.5) for $\Delta_{n,j+1}$ and $Y_{n,j+1}$, respectively, to compute

$$\Delta_{n,j+1} = \alpha_n Y_{n,j+1} - Y_{n,j+1} \beta_j + Y_{n-1,j+1} - Y_{n,j} = \alpha_n \mu_{j+1} - \mu_{j+1} \beta_{j+1} = 0, \quad (5.17)$$

where the last equality is the consequence of (5.2). Due to induction hypothesis and (5.17), we may apply implication (5.11) to conclude that $\Delta_{k,i} = 0$ also for all k, i such that $k + i = r + 1$. By induction principle, $\Delta_{k,j} = 0$ for all $k = 1, \dots, n$, $j = 1, \dots, m$ and since $\Delta = [\Delta_{k,j}] = AY - YB$ (by definition (5.9)), we conclude that Y is a solution to (5.1).

It remains to show that any Y solving (5.1) is necessarily of the form (5.2)–(5.5). By Theorem 4.5 (part (2)), the entries $Y_{k,j}$ (for $k + j \leq n$) are the same for any solution Y to the equation (4.50). Since the zero matrix is a solution, it then follows that $Y_{k,j} = 0$ whenever $k + j \leq n$. Equating the entries from the bottom row in (5.1), we get, on account of (4.1),

$$\alpha_n Y_{n,1} - Y_{n,1} \beta_1 = 0, \quad (5.18)$$

$$\alpha_n Y_{n,j} - Y_{n,j} \beta_j + Y_{n-1,j} - Y_{n,j-1} = 0 \quad (j = 2, \dots, m). \quad (5.19)$$

Equation (5.18) simply means that $Y_{n,1} = \mu u_1$ can be picked arbitrarily in Π_{α_n, β_1} , the solution set of the scalar Sylvester equation (5.18). By (the proof of) Theorem 4.5, any solution Y to the equation (5.1) also satisfies

$$\tilde{A}Y - Y\tilde{B} = 0, \quad (5.20)$$

where \tilde{A} and \tilde{B} are given in (3.10) and (4.39). Equating the $(n-1, 2)$ entries in (5.20) we get $(\bar{\alpha}_n - \alpha_{n-1})Y_{n-1,2} - Y_{n,1}(\bar{\beta}_2 - \beta_1) = 0$, which is equivalent to (5.7) and thus, verifies (5.4) for $k = n - 1$ and $j = 2$. Then we can apply formula (5.14) for $k = n - 1$ and $j = 1$ to get (we recall that $Y_{n-2,2} = Y_{n-1,1} = 0$)

$$Y_{n-1,2} - Y_{n,1} = \alpha_n (\bar{\alpha}_n - \alpha_{n-1})^{-1} Y_{n,1} - (\bar{\alpha}_n - \alpha_{n-1}) Y_{n,1} \beta_2,$$

which, being substituted into equation (5.19) (for $j = 1$) leads us to

$$0 = \alpha_n Y_{n,2} - Y_{n,2} \beta_2 + Y_{n-1,2} - Y_{n,1} \\ = \alpha_n (Y_{n,2} + (\bar{\alpha}_n - \alpha_{n-1})^{-1} Y_{n,1}) - (Y_{n,2} + (\bar{\alpha}_n - \alpha_{n-1})^{-1} Y_{n,1}) \beta_2.$$

The latter equality tells us that the element

$$\mu_2 := Y_{n,2} + (\bar{\alpha}_n - \alpha_{n-1})^{-1} Y_{n,1}$$

must satisfy equation (5.2) for $j = 2$, which verifies formula (5.5) for $j = 1$. So far, we verified formulas (5.4) and (5.5) for the entries from the two leftmost columns

of Y . Let us assume the formulas hold true for all entries from the $r \geq 2$ leftmost columns and let us write explicitly conditions $[\tilde{A}Y - Y\tilde{B}]_{k+1,r+1} = 0$ (see (5.20)) for $k = 1, \dots, n-1$; in view of (3.10) and (4.39), we have

$$(\alpha_k - \bar{\alpha}_{k+1})Y_{k,r+1} + Y_{k-1,r+1} - Y_{k+1,r}(\beta_r - \bar{\beta}_{r+1}) - Y_{k+1,r-1} = 0$$

which being solved for $Y_{k,r+1}$ verifies formulas (5.4) for $j = r+1$ and $k \leq n-1$. Since the formulas (5.4) and (5.5) hold for $j = r$ and all $k \leq n$ by the assumption, equality (5.16) holds for $j = r$. Substituting this equality into equation (5.19) (for $j = r+1$) gives

$$\begin{aligned} 0 &= \alpha_n Y_{n,r+1} - Y_{n,r+1} \beta_{r+1} + Y_{n-1,r+1} - Y_{n,r} \\ &= \alpha_n \left(Y_{n,r+1} + \sum_{\ell=1}^r \left(\prod_{i=1}^{\ell} (\bar{\alpha}_{n-i+1} - \alpha_{n-i})^{-1} \right) Y_{n-\ell+1,r} \right) \\ &\quad - \left(Y_{n,r+1} + \sum_{\ell=1}^r \left(\prod_{i=1}^{\ell} (\bar{\alpha}_{n-i+1} - \alpha_{n-i})^{-1} \right) Y_{n-\ell+1,r} \right) \beta_{r+1}. \end{aligned}$$

We conclude that the element

$$\mu_{r+1} := Y_{n,r+1} + \sum_{\ell=1}^r \left(\prod_{i=1}^{\ell} (\bar{\alpha}_{n-i+1} - \alpha_{n-i})^{-1} \right) Y_{n-\ell+1,r}$$

must be in $\Pi_{\alpha_n, \beta_{r+1}}$ which verifies (5.5) for $j = r+1$. By induction principle, formulas (5.4) and (5.5) hold for all $j = 1, \dots, m$. \square

Corollary 5.2. *Let $A = \mathcal{I}_n(\alpha)$ and $B = \mathcal{I}_m(\beta)$ ($n \geq m$) be Jordan blocks based on the elements $\alpha \sim \beta$. A matrix $Y = [Y_{k,j}] \in \mathbb{H}^{n \times m}$ satisfies $AY = YB$ if and only if*

$$Y_{k,j} = 0 \quad (2 \leq k+j \leq n) \quad \text{and} \quad Y_{k,j} = \mu_{k+j-n} \quad (n+1 \leq k+j \leq n+m), \quad (5.21)$$

where $\mu_1, \dots, \mu_m \in \mathbb{H}$ are any elements from the plane $\Pi_{\alpha, \beta}$ (see (4.4) and (4.5)), i.e.,

$$\alpha \mu_i = \mu_i \beta \quad \text{for } i = 1, \dots, m. \quad (5.22)$$

Proof. It suffices to let $\alpha_k = \alpha$ and $\beta_j = \beta$ in Theorem 5.1 and to show that formulas (5.3)–(5.5) amount to (5.21) in this particular setting. This is clearly true for $k+1 \leq n$ and for $Y_{n,1} = \mu_1$. For the rest, we will use equalities

$$\bar{\alpha} \mu_i = \mu_i \bar{\beta} \quad (i = 1, \dots, m), \quad (5.23)$$

which follow from (5.22) since $\text{Re}(\alpha) = \text{Re}(\beta)$, by (2.1). Due to equalities (5.22) and (5.23), formula (5.6) takes the form

$$Y_{n-j,j+1} = (\bar{\alpha} - \alpha)^{-j} \mu_1 (\bar{\beta} - \beta)^j = \mu_1 \quad (j = 1, \dots, m-1).$$

By (5.8), $Y_{n,2}$ is necessarily of the form $Y_{n,2} = \mu_2 - (\bar{\alpha} - \alpha)^{-1} \mu_1$ where μ_2 is any element in $\Pi_{\alpha, \beta}$. But since

$$\alpha (\bar{\alpha} - \alpha)^{-1} \mu_1 = (\bar{\alpha} - \alpha)^{-1} \alpha (\bar{\alpha} - \alpha)^{-1} \mu_1 \beta,$$

i.e., since $(\bar{\alpha} - \alpha)^{-1}\mu_1 \in \Pi_{\alpha,\beta}$, it follows that $Y_{n,2}$ can be chosen arbitrarily in $\Pi_{\alpha,\beta}$. In other words, we may let $Y_{n,2} = \mu_2 \in \Pi_{\alpha,\beta}$. Since $Y_{n-j,j+1} = \mu_1$ for all $j = 0, \dots, m-1$, the formula (5.4) (for $k+j = n+2$) now gives

$$Y_{n-j+1,j+1} = (\bar{\alpha} - \alpha)^{-j}\mu_2(\bar{\beta} - \beta)^j = \mu_2 \quad (j = 2, \dots, m-1)$$

verifying the Hankel structure of $\mathcal{D}_{n+1}(Y)$. Now we have by (5.5),

$$Y_{n,3} = \mu_3 - (\bar{\alpha} - \alpha)^{-1}\mu_2 - (\bar{\alpha} - \alpha)^{-2}\mu_1,$$

and since the second and the third terms on the right belong to $\Pi_{\alpha,\beta}$ and μ_3 can be chosen in $\Pi_{\alpha,\beta}$ arbitrarily, we may let $Y_{n,3} = \mu_3 \in \Pi_{\alpha,\beta}$. Now we use formula (5.4) and take into account the Hankel structure of $\mathcal{D}_{n+1}(Y)$ to verify the Hankel structure of $\mathcal{D}_{n+2}(Y)$. Repeating this argument m times we arrive at the desired conclusion. \square

Recall that all results in Sections 4 and 5 were obtained under assumption that $n \geq m$. The assumption is not restrictive in the following sense: if $m \geq n$, then one can apply the obtained results to the adjoint equation (3.24) as in the proof of Theorem 3.9, i.e., to replace A, B and C by B^*, A^* and C^* , respectively. We omit further details.

REFERENCES

1. V. Bolotnikov, *Polynomial interpolation over quaternions*, J. Math. Anal. Appl. **421** (2015), no. 1, 567–590.
2. V. Bolotnikov, *Zeros and factorizations of quaternion polynomials: the algorithmic approach*, Preprint.
3. V. Bolotnikov, *Quasi-ideals in the ring of quaternion polynomials*, Preprint.
4. J. L. Brenner, *Matrices of quaternions*, Pacific J. Math., **1** (1951), 329–335.
5. F. Cecioni, *Sulle equazioni fra matrici $AX = XB$, $X^n = A$* , R. Acad. dei Lincei, Rend. **18** (1909), 566–571.
6. F. Cecioni, *Sopra alcune operazioni algebriche sulle matrici*, Ann. R. Scuola Norm. Sup. Pisa **11** (1910), 1–40.
7. G. Frobenius, *Über die mit einer Matrix vertauschbaren Matrizen*, Sitzungsber. Preuss. Akad. Wiss., 1910, 3–15.
8. I. Gohberg, P. Lancaster, and L. Rodman, *Matrix polynomials*, Academic Press, , New York-London, 1982.
9. L. Huang, *The matrix equation $AXB - CXD = E$ over the quaternion field*, Linear Algebra Appl. **234** (1996), 197–208.
10. A. Jameson, *Solutions of the equation $AX + XB = C$ by inverse of an $M \times M$ or $N \times N$ matrix*, SIAM J. Appl. Math. **16** (1968) 1021–1022.
11. P. Lancaster, *Explicit solutions of linear matrix equations*, SIAM Rev. **12** 1970 544–566.
12. H. C. Lee, *Eigenvalues and canonical forms of matrices with quaternion coefficients*, Proc. Roy. Irish Acad. **52** (1949), 253–260.
13. E. C. Ma, *A finite series solution of the matrix equation $AX - XB = C$* , SIAM J. Appl. Math. **14** (1966) 490–495.
14. M. Rosenblum, *On the operator equation $BX - XA = Q$* , Duke Math. J. **23** (1956) 263270.
15. D. E. Rutherford, *On the solution of the matrix equation $AX + XB = C$* , Proc. Akad. Wet. Amsterdam **35** (1932), 54–59.
16. C. Song, and G. Chen, *On solutions of matrix equation $XF - AX = C$ and $XF - A\tilde{X} = C$ over quaternion fie*, J. Appl. Math. Comput. **37** (2011), no. 1-2, 57–68.
17. J. J. Sylvester, *Sur l'equation en matrices $px = xq$* , C. R. Acad. Sci. Paris **99** (1884) 67–71, 115–116.

18. P. G. Tait, *An elementary treatise on quaternions*, Oxford, Clarendon Press, 1867.
19. J.H.M. Wedderburn, *On division algebras*, Trans. Amer. Math. Soc. **22** (1921), 129–135.
20. N.Wiegmann, *Some theorems on matrices with real quaternion elements*, Canad. J. Math. **7** (1955) 191–201.

WILLIAMSBURG